# Malliavin calculus for marked binomial processes: portfolio optimisation in the trinomial model and compound Poisson approximation

Hélène Halconruy*

**Abstract**

In this paper we develop a stochastic analysis for marked binomial processes, that can be viewed as the discrete analogues of marked Poisson processes. The starting point is the statement of a chaotic expansion for square-integrable (marked binomial) functionals, prior to the elaboration of a Markov-Malliavin structure within this framework. We take advantage of the new formalism to deal with two main applications. First, we revisit the Chen-Stein method for the (compound) Poisson approximation which we perform in the paradigm of the built Markov-Malliavin structure, before studying in the second one the problem of portfolio optimisation in the trinomial model.

**Keywords:** Marked binomial process, chaotic decomposition, Mehler's formula, Mallliavin calculus, Gamma calculus, Stein's method, discrete market model, portfolio optimisation, compound Poisson approximation.

## 1 Introduction

This paper is motivated by two applications: the compound Poisson approximation by a revisited Chen-Stein method and a problem of portfolio optimisation in the trinomial model. Apparently unrelated, they are both possible by-products of Malliavin calculus. The eponymous theory designed by Paul Malliavin in the 70's was initially elaborated to provide an infinite dimensional differential calculus for the Wiener space, and further extended to other settings. Thus one can find in the literature stochastic variational calculus for Gaussian processes in general (see Nualart [48], Nourdin and Peccati [45]), Poisson processes (see Bichteler *et al.* [10] for a variational approach, Nualart and Vives [51] or Privault [55] for a chaotic approach), Lévy processes (see Nualart and Schoutens [50]), Rademacher processes (see Privault [56]), and more recently for independent random variables (developed independently by Duerinckx, Gloria and Otto in [24], Decreusefond and Halconruy in [20]). Even if the multiplicity of approaches and the variety of canonical spaces on which Malliavin calculus operates appeared at first glance to be an obstacle to a complete unifying theory, one can find nevertheless a common terminology for all these formalisms around the notions of Malliavin operators (gradient D, divergence $\delta$, Ornstein-Uhlenbeck operator L and semi-group $(\mathrm{P}_t)_{t\in\mathbf{R}_+}$) and the fundamental relationship between the gradient operator and divergence (defined as the gradient adjoint): the integration by parts formula. In fact, this is part of a deeper structure in which

*Mathematics Research Unit, Luxembourg University, Luxembourg. e-mail: helene.halconruy@uni.lu

the L operator would be, by virtue of its dual role, the cornerstone: as the Ornstein-Uhlenbeck operator, it generates an underlying Markovian structure, and, as the Laplacian operator, it satisfies $L = -\delta D$. This raises, via the integration by parts formula, a Dirichlet form $\mathcal{E}$, so that any Malliavin equipment comes with a Dirichlet structure (see Decreusefond [18]). Besides, the form $\mathcal{E}$ appears as the energy function associated to the *carré du champ* operator $\Gamma$. The naturally emerging Gamma-Malliavin structure provides an ideal framework for future applications. Indeed the operator $\Gamma$ is one efficient to deduce quantitative limits by means of...

... *Stein's method*, that frames the first application we have investigated. Initially designed to quantify the errors in the normal approximation by sums of random variables having a stationary dependence structure, Stein's method stood out as one (not to say the one) efficient way to derive distance bounds between two probability measures (referred to as initial distribution and target distribution) with respect to a certain probabilistic metrics. It seems to be split into two stages; the first is to take advantage of the target law characterisation to convert the original problem into that of bounding some functional of the initial space, whereas the second one aims at developing tools to tackle with this new expression. This latter can be performed by using exchangeable pairs (see Barbour and Chen [5]) or other forms of couplings such as zero- or size-biased couplings. In a path-breaking work, Nourdin and Peccati (see [44],[45]) showed that the transformation step can be advantageously made simple using integration by parts in the sense of Malliavin calculus, and by the same gave an intersection to the two theories.
This approach is efficient provided there exists a Malliavin gradient on the initial space. This was handled to derive bounds for normal approximation by functionals acting on spaces equipped with a Malliavin structure such as Wiener chaoses (see Nourdin and Peccati [43], with Reinert [46], [49], Víquez [73]), Poisson functionals (Lachièze-Rey and Peccati [34], Peccati *et al.* [54], Schulte [67]), functionals of Rademacher (see Reinert *et al.* [47], Zheng [75]). On the other hand, Stein's method was developed to other target distributions; its most famous declination is undoubtedly the *Chen-Stein* method that deals with (compound) Poisson approximation (see Chen [13]). This was first performed using location arguments by the introduction of "neighbourhood of dependence" sets to deal with Poisson approximation (see Arratia *et al* [2] and [1], Barbour *et al* [8]) or compound Poisson one (see Barbour *et al* [6], [9]). Instigated first by Peccati in [53], the Chen-Stein method was also combined to Malliavin calculus to provide Poisson approximation bounds for point processes (see Torrisi [72]), Rademacher functionals (Krokowski [33], Privault and Torrisi [59]), or multiple integrals (see Bourguin and Peccati [11]). Last, as mentioned above, the Stein-Malliavin method was recently improved by exploiting the underlying Markovian structure (and in particular by using the carré du champ operator) to get quantitative limits by overcoming possible combinatorial difficulties that may arise from the use of multiplication formulae on configuration spaces. Finding its theoretical roots in the innovative works of Azmoodeh, Campese and Poly [3], Azmoodeh *et al.* [4], Ledoux [39], this approach successfully succeeds exploiting the powerful techniques of Markov generators within Gamma-calculus.
Our first application takes place in this landscape. We propose a new point of view to address the problem of Poisson (respectively compound Poisson) approximation for the longest perfect head run in a coin tossing experiment (respectively for the occurrence of a rare word in a DNA sequence) within the Chen-Stein method. Indeed, these were treated so far by means of identified "dependent neighbours"; here, we handle these problems through a Stein-Markov-Malliavin method based on the keystone operator L (we cannot really take advantage of the

operator carré du champ within our framework) for marked binomial processes.

The second motivation to elaborate a stochastic variational calculus for marked binomial processes, comes from one famous scope of Malliavin calculus: finance. One of the most successful area of financial mathematics deals with option pricing and more generally valuation of contingent contracts. A claim is a non-negative random variable F often assumed to be square-integrable that models the *payoff* (the value of the option at expiry) at a fixed-term maturity $T$ of some asset. The simple example is given by a European call (resp. put) option based on the asset S with expiration date $T$ and strike price $K$, defined by $\mathrm{F} = (\mathrm{S}_T - K)_+$ (respectively $\mathrm{F} = (K - \mathrm{S}_T)_+$). The claim is said to be *attainable* or *duplicable* if there exists a self-financing portfolio of value F at expiry, and *redundant* if this replication is only based on the existing assets. In *complete* markets, all claims are reachable. The Cox-Ross-Rubinstein (CRR for short) model is the simplest discrete example of complete market (see Cox, Ross and Rubinstein [15], Rendleman and Bartter [61]). The so-called Fundamental Asset Pricing Theorem (FAPT) asserts that an arbitrage-free market is complete if and only if there exists a unique probability measure, equivalent to the initial under which the discounted price process is a martingale (in a discrete time setting see Jacod [32], Schachermayer [65]). The *prime* or *fair price*, equal to the initial value of the replicating portfolio and more generally to its value at any time can be written as the (conditional) expectation with respect to this unique *martingale measure* (for the seminal paper see Harrison and Kreps [40]). Besides, an explicit formula of the replicating strategy in terms of Malliavin derivative can be provided by the application of Clark-Ocone formula in the CRR model (see chapter one in Privault [56] and [57]). Several approaches were developed to address the problem of hedging in an incomplete market, as the trinomial model we are focusing on in this paper. To name but a few, one consists in "completing" the market by introducing new securities in an equilibrium approach (see Hakansson [30], Boyle and Tan [12]). Within the no-arbitrage framework, another approach consists in exhibiting the so-called *minimal martingale measure* from the set of equivalent martingale measures. It is related to risk-minimizing strategies (Föllmer and Sondermann [27], Schweizer [68], [70]) or portfolio expected utility maximization under constraints (Delbaen and Schachermayer [22], Frittelli [28], Runggaldier [64]). In the application we choose to develop, we are less concerned by stating a valuation formula than to determine the optimized portfolio composition of minimal risk (in a sense to be defined) for an a priori non-attainable claim. When the claim is reachable, we get then a hedging formula, that has not been done - to our knowledge - in the trinomial frame.
Our initial aim was to transpose the criteria stated by Föllmer and Sondermann in [27] into the frame of the trinomial model (underpinned by sequences of $\{-1, 0, 1\}$-valued independent random variables) in order to determine the less risky approximating portfolio and derive the explicit expression of the corresponding optimizing strategy from the Clark formula stated for independent random variables (see Decreusefond and Halconruy [20] Theorem 3.3). The lack of a martingale representation theorem in that latter framework, made it impossible to derive a hedging formula from Clark's (as done in binomial or Black-Scholes model); this led us to replace the trinomial model with what we called *a ternary model*, also composed of two assets (a riskless and a risky asset) and subtended by a marked binomial process where the mark space consists of two elements that indicate whether the price sequence (of the risky asset) rises or falls. Besides, if the probabilities of the occurrence of jumps and marks are properly defined, the ternary model is in fact equivalent in law to the trinomial model, so that

all results "on expectation" hold in this latter.

The interest of this work is therefore twofold; from a theoretical point of view, is built up there a variational calculus for binomial processes within a unifying Markov-Malliavin structure including pre-existing theories (in the Wiener or Poisson space). From an application angle, this new formalism is prone to offer an alternative point of view on Chen-Stein method and to produce an explicit portfolio optimisation formula in the trinomial model that had - to the best of our knowledge - not been done (in that way) so far for multiple periods.

The paper is structured as follows. In section 2 we give some elements of stochastic analysis for marked binomial processes and state a chaotic expansion result for any square-integrable binomial functional. From the successive development of a variational calculus in $\mathrm{L}^1$ and $\mathrm{L}^2$-contexts, we formalize a Markov-Malliavin structure for marked binomial processes in Section 3 whereas the Section 4 is devoted to the statement of a Girsanov theorem and a Clark formula within this framework.The Section 5 is dedicated to the two main applications of our formalism: the (compound) Poisson approximation by Chen-Stein method and the portfolio optimisation in the trinomial model. Almost all proofs are postponed to Section 6.

# 2 Stochastic analysis for marked binomial processes, part I

The first part of this present section is devoted to the introduction of notation and the main object of interest, the *marked binomial process*. Throughout, $(\Omega, \mathcal{A}, \mathbf{P})$ will be an abstract probability space assumed to be wide enough to support all random objects in question.

## 2.1 Framework

Consider the measurable space $(\mathbb{X}, \mathcal{X})$ where $\mathbb{X} = \mathbf{N} \times \mathrm{E}$, and $(\mathrm{E}, \mathcal{B}(\mathrm{E}))$, the *mark space*, is a Borel space. Without any other indication, E will designate a countable (possibly finite) subset of $\mathbf{Z}$. Nevertheless, our construction may be extended to any subset of $\mathbf{R}$ and we provide later additional elements to formalize it. Denote by $\mathfrak{N}_{\mathbb{X}}$ (respectively $\widehat{\mathfrak{N}}_{\mathbb{X}}$) the space of simple, integer-valued, $\sigma$-finite (respectively finite) measures on $\mathbb{X}$. Let $\mathcal{N}^{\mathbb{X}}$ be the smallest $\sigma$-field of subsets of $\mathfrak{N}_{\mathbb{X}}$ such that the mapping $\chi \in \mathfrak{N}_{\mathbb{X}} \mapsto \chi(\mathrm{A})$ is measurable for all $\mathrm{A} \in \mathcal{X}$. A *point process* (respectively finite point process) - or random counting measure - is a random element $\eta$ in $\mathfrak{N}_{\mathbb{X}}$ (respectively in $\widehat{\mathfrak{N}}_{\mathbb{X}}$) that satisfies $\eta(\mathrm{A}) \in \mathbf{Z}_+ \cup \{\infty\}$ for all $\mathrm{A} \in \mathcal{X}$. In its very definition $(\mathrm{E}, \mathcal{B}(\mathrm{E}))$ is a very simple Polish space endowed with its Borel $\sigma$-field so that we may and will assume that any element $\eta$ of $\mathfrak{N}_{\mathbb{X}}$ is *proper*, i.e. can be $\mathbf{P}$-a.s. written as

$$\eta = \sum_{n=1}^{\eta(\mathbb{X})} \delta_{\mathrm{X}_n}, \tag{2.1}$$

where $\{\mathrm{X}_n,\, n \geqslant 1\}$ denotes a countable collection of $\mathbb{X}$-valued random elements, and for $x \in \mathbb{X}$, $\delta_x$ is the Dirac measure at $x$. For a complete exposé on the subject of point processes, the reader can refer to the monograph of Last and Penrose ([38], section 6.1) or Last [35] from that our presentation is largely inspired by. A *binomial marked process* is a particular point process $\eta$ defined as follows; let $\lambda \in (0,1)$, and consider a Bernoulli process (see for instance Decreusefond and Moyal [21], definition 6.6) of parameter $\lambda$, described by a sequence of jump times $(\mathrm{T}_t)_{t\in\mathbf{Z}_+}$, such that for any $t \in \mathbf{Z}_+$, the $t$-th arrival time $\mathrm{T}_t$ is defined by $\mathrm{T}_0 = 0$ and $\mathrm{T}_t = \sum_{s=1}^{t} \xi_s$, and where the inter-arrival variables $\{\xi_t,\, t \in \mathbf{N}\}$ are independent and identically distributed by a geometric law of parameter $\lambda$. In analogy with marked Poisson processes (see Last and Penrose [38], chapter 7), we can set that $\eta$ is $\mathbf{P}$-a.s. represented as

$$\eta = \sum_{t=1}^{\infty} \mathbf{1}_{\{\mathrm{T}_t<\infty\}} \delta_{(\mathrm{T}_t,\mathrm{V}_t)}, \tag{2.2}$$

where $\{\mathrm{V}_t,\, t \in \mathbf{N}\}$ is a collection of E-valued random elements such that almost surely $\eta(\mathrm{T}_t, \mathrm{V}_t) = 1$, for $\mathrm{T}_t < \infty$, and that are independent of the underlying jump process $\mathrm{N} = (\mathrm{N}_t)_{t\in\mathbf{Z}_+}$ defined by $\mathrm{N}_0 = 0$ and $\mathrm{N}_t = \sum_{s\in\mathbf{N}} \mathbf{1}_{\{\mathrm{T}_s\leqslant t\}}$. By a slight abuse of notation, we shall write $(t, k) \in \eta$ in order to indicate that the point $(t, k) \in \mathbb{X}$ is charged by the random measure $\eta$. Note that for any $t \in \mathbf{N}$, $\mathrm{N}_t$ is a binomial random variable of mean $\lambda t$.
We may and will assume that $\mathcal{A} = \sigma(\eta) =: \mathcal{F}$ where $\mathcal{F} = (\mathcal{F}_t)_{t\in\mathbf{N}}$ is the canonical filtration defined from $\eta$ by

$$\mathcal{F}_0 := \{\emptyset, \Omega\} \quad \text{and} \quad \mathcal{F}_t := \sigma\left\{ \sum_{(s,k)} \eta(s, k),\ s \leqslant t,\ k \in \mathrm{E} \right\}.$$

Let $\mathbf{Q}$ be the common distribution of the $\mathrm{V}_t$ and $\mathbf{P}_\eta = \mathbf{P} \circ \eta^{-1}$ be the image measure of $\mathbf{P}$ under $\eta$ on the space $(\mathfrak{N}_{\mathbb{X}}, \mathcal{N}^{\mathbb{X}})$ i.e. the distribution of $\eta$; its compensator - the intensity of $\eta$ - is the measure $\nu$ defined on $\mathcal{X}$ by

$$\nu(\mathrm{A}) = \sum_{(t,k)\in\mathrm{A}} \sum_{s\in\mathbf{N}} \Big( \lambda\delta_s(\{t\}) \otimes \sum_{\ell\in\mathrm{E}} \mathbf{Q}(\{\ell\})\delta_\ell(\{k\}) \Big)\ ;\ \mathrm{A} \in \mathcal{X}.$$

Throughout, we denote by $\mathbf{R}(\mathfrak{N}_{\mathbb{X}})$ the class of real-valued measurable functions $\mathfrak{f}$ on $\mathfrak{N}_{\mathbb{X}}$ and by $\mathcal{L}^0(\Omega) := \mathcal{L}^0(\Omega, \mathcal{A})$ the class of real-valued measurable functions F on $\Omega$. Since $\mathcal{A} = \sigma(\eta)$, for any $\mathrm{F} \in \mathcal{L}^0(\Omega)$, there exists a function $\mathfrak{f} \in \mathbf{R}(\mathfrak{N}_{\mathbb{X}})$ such that $\mathrm{F} = \mathfrak{f}(\eta)$. The function $\mathfrak{f}$ is called a *representative* of F and is $\mathbf{P}\otimes\eta^{-1}$-a.s. uniquely defined. By default, the representative of a random variable $\mathrm{F} \in \mathcal{L}^0(\Omega)$ will be noted by the corresponding gothic lowercase letter, $\mathfrak{f}$.
Last, for $p \in \mathbf{N}$, we define $\mathrm{L}^p(\mathbf{P}) := \mathrm{L}^p(\Omega, \mathcal{A}, \mathbf{P})$ the set of $p$-integrable functions on $\Omega$ with respect to $\mathbf{P}$.

**Remark 2.1.** The marked binomial process $\eta$ can be equivalently written as $\sum_{t\in\mathbf{N}} \delta_{(\Delta\mathrm{N}_t,\mathrm{W}_t)}$, where for any $t \in \mathbf{N}$ the random variables $\Delta\mathrm{N}_t$ and $\mathrm{W}_t$ are defined by

$$\Delta\mathrm{N}_t = \mathrm{N}_t - \mathrm{N}_{t-1} = \sum_{k\in\mathrm{E}} \mathbf{1}_{\{(t,k)\in\eta\}} \quad \text{and} \quad \mathrm{W}_t = \sum_{k\in\mathrm{E}} k\mathbf{1}_{\{(t,k)\in\eta\}}.$$

The variables $\Delta\mathrm{N}_t$ and $\mathrm{W}_t$ thus defined play a major part and we will often refer to them. Indeed, $\Delta\mathrm{N}_t$ indicates whether there is a jump at time $t$, and, if so, the variable $\mathrm{W}_t$ gives its corresponding mark $k$.

If $(\mathrm{V}_t)_{t\in\mathbf{N}}$ is a sequence of independent $\mathbf{R}$-valued random variables with common distribution $\mathbf{Q}$ and that are independent of the process N, we can define the *compound binomial process* $\mathrm{Y} = (\mathrm{Y}_t)_{t\in\mathbf{N}}$ of intensity $\nu$ by

$$\mathrm{Y}_t := \sum_{s=1}^{\mathrm{N}_t} \mathrm{V}_s. \tag{2.3}$$

The corresponding compensated process denoted $\overline{\mathrm{Y}} = (\overline{\mathrm{Y}}_t)_{t\in\mathbf{N}}$ defined by

$$\overline{\mathrm{Y}}_t := \Big(\sum_{s=1}^{\mathrm{N}_t} \mathrm{V}_s\Big) - \lambda \mathbf{Q}(\{k\})t \; ; \; t \in \mathbf{N}, \tag{2.4}$$

is a $\mathcal{F}$-martingale. In their very definitions, $\eta$, Y and $\overline{\mathrm{Y}}$ are the discrete analogues of the marked, compound Poisson and compensated compound Poisson processes.

## 2.2 Integration with respect to a binomial marked process

A process $u = (u_{(t,k)})_{(t,k)\in\mathbb{X}}$ is a measurable random variable defined on $(\mathfrak{N}(\mathbb{X}) \times \mathbb{X}, \mathcal{F} \otimes \mathcal{X})$ that can be written $u = \sum_{(t,k)\in\mathbb{X}} \mathfrak{u}(\eta,(t,k))\mathbf{1}_{(t,k)}$, where $\{\mathfrak{u}(\eta,(t,k)),\, (t,k) \in \mathbb{X}\}$ is a family of measurable functions from $\mathfrak{N}_{\mathbb{X}} \times \mathbb{X}$ to $\mathbf{R}$ and $\mathfrak{u}$ is called the *representative* of $u$. As for random variables, the representative of a process will be noted by a Gothic letter. For instance, considering a process $r$, its representative will be denoted by $\mathfrak{r}$. The following assumption holds throughout this subsection.

**Assumption 2.2.** *There exists a discrete-time process* $\mathrm{R} = (\mathrm{R}_{(t,k)},\, t \in \mathbf{Z}_+, k \in \mathrm{E})$ *where* $\mathrm{R}_{(t,k)} = \mathfrak{r}(\eta,(t,k))$ *is defined on* $(\Omega, \mathcal{F}, \mathbf{P})$ *and that satisfies the following hypotheses:*

1. *The process* $\Big(\sum_{k\in\mathrm{E}} \mathrm{R}_{(t,k)}\Big)_{t\in\mathbf{Z}_+}$ *is a* $\mathcal{F}$*-martingale,*
2. *The family* $\mathcal{R} = \{\Delta\mathrm{R}_{(t,k)},\, t \in \mathbf{N},\, k \in \mathrm{E}\}$ *is orthogonal for the scalar product* $(\mathrm{X},\mathrm{Y}) \in \mathrm{L}^2(\mathbf{P}) \mapsto \mathbf{E}[\mathrm{XY}]$ *and* $\Delta\mathrm{R}_{(t,k)}$ *and* $\Delta\mathrm{R}_{(s,k)}$ *are identically distributed for all* $t, s \in \mathbf{N}$, $k \in \mathrm{E}$. *We denote* $\mathbf{E}\big[(\Delta\mathrm{R}_{(t,k)})^2\big] =: \kappa_k$ *for any* $(t,k) \in \mathbb{X}$.

### 2.2.1 Stochastic integrals

Throughout the paper, we adopt the following set notations; we denote $\{a : b\} := \{a, \dots, b\}$ for any $a, b \in \mathbf{Z}$ such that $a < b$, and $\mathbb{X}_t := \{1 : t\} \times \mathrm{E}$ for any $t \in \mathbf{N}$. By convention, $\{1 : 0\} = \emptyset$. Any $n$-tuple of $\mathbb{X}^n$ can be denoted by bold letters; for instance, $(\mathbf{t}_n, \mathbf{k}_n) = \big((t_1,k_1), \cdots, (t_n,k_n)\big)$. For any $\mathrm{A} \in \mathcal{X}$, we denote $\mathrm{A}^{n,<} = \{(\mathbf{t}_n,\mathbf{k}_n) \in \mathbb{X}^n : t_1 < \cdots < t_n\}$, the corresponding time-ordered set, and $\mathrm{A}^{n,\neq} = \{(\mathbf{t}_n,\mathbf{k}_n) \in \mathrm{A}^n : \forall i \neq j,\, t_i \neq t_j\}$, the set with pairwise distinct (in time) elements.

We denote by $\mathrm{L}^2(\mathbf{P}\otimes\nu)$ the Hilbert space of processes that are square-integrable with respect to the measure $\mathbf{P}\otimes\nu$, for which we define the corresponding inner product and norm by

$$\langle u, v\rangle_{\mathrm{L}^2(\mathbf{P}\otimes\nu)} = \mathbf{E}\Big[\int_{\mathbb{X}} \mathfrak{u}(\eta,(t,k))\mathfrak{v}(\eta,(t,k))\,\mathrm{d}\nu(t,k)\Big] \text{ and } \|u\|^2_{\mathrm{L}^2(\mathbf{P}\otimes\nu)} = \mathbf{E}\Big[\int_{\mathbb{X}} \mathfrak{u}(\eta,(t,k))^2\,\mathrm{d}\nu(t,k)\Big].$$

**Definition 2.3.** The set of *simple processes*, denoted by $\mathcal{U}$ is the set of random variables of the form

$$u = \sum_{(t,k)\in\mathbb{X}_T} \mathfrak{u}(\eta,(t,k))\mathbf{1}_{(t,k)}, \tag{2.5}$$

where $T \in \mathbf{N}$, and $\mathfrak{u}$ is the representative of $u$. Let $\mathcal{P}$ denote the subspace of $\mathcal{U}$ made of simple predictable processes i.e. of the form (2.5) where $\mathfrak{u}(\eta,(t,\cdot))$ is $\mathcal{F}_{t-1}$-measurable for any $t \in \{1 : T\}$.

**Proposition 2.4.** *Any process $u \in \mathcal{U}$ of representative $\mathfrak{u}$ is integrable with respect to the process* R *by the formula*

$$\mathrm{J}_1(u\,;\mathcal{R}) = \sum_{(t,k)\in\mathbb{X}} \mathfrak{u}(\eta,(t,k))\Delta\mathrm{R}_{(t,k)}.$$

*The so-called $\mathcal{R}$-stochastic integral $\mathrm{J}_1(u\,;\,\mathcal{R})$ of $u$ extends to square-integrable predictable processes via the (conditional) isometry formula*

$$\mathbf{E}\left[\left|\mathrm{J}_1\big(\mathbf{1}_{[t,\infty)}u\,;\mathcal{R}\big)\right|^2 \,|\, \mathcal{F}_{t-1}\right] = \mathbf{E}\left[\left\|\mathbf{1}_{[t,\infty)}u\right\|^2_{\mathrm{L}^2(\mathbb{X},\tilde{\nu})} \,|\, \mathcal{F}_{t-1}\right], \tag{2.6}$$

*where $\tilde{\nu}$ is the measure on $\mathbb{X}$ defined by $\tilde{\nu}(\{(t,k)\}) = \kappa_k\nu(\{(t,k)\}))$, for any $(t,k)\in\mathbb{X}$.*

### 2.2.2 Multiple integrals

In order to define (multiple) stochastic integrals, we work in a space of symmetrical functions. Our construction follows closely that depicted by Privault (see [56], chapter 6); in a certain sense we transpose it into our context. The space $\mathrm{L}^2(\mathbb{X},\nu)^{\circ 0}$ is by convention identified to $\mathbf{R}$; let thus for any $f\in\mathrm{L}^2(\mathbb{X},\nu)^{\circ 0}$, $\mathrm{J}_0(f_0) = f_0$.

**Definition 2.5.** For $n\in\mathbf{N}$, let $\mathrm{L}^2(\mathbb{X},\nu)^{\circ n}$ denote the subspace of $\mathrm{L}^2(\mathbb{X},\nu)^{\otimes n} = \mathrm{L}^2(\mathbb{X},\nu)^n$ composed of the functions $f_n\in\mathbf{R}(\mathbb{X}^n)$ symmetric in their $n$ variables, i.e. such that for any permutation $\tau$ of $\{1,\dots,n\}$, $f_n\big((t_{\tau(1)},k_{\tau(1)}),\cdots,(t_{\tau(n)},k_{\tau(n)})\big) = f_n\big((t_1,k_1),\cdots,(t_n,k_n)\big)$, for all $(t_1,k_1),\dots,(t_n,k_n)\in\mathbb{X}$. The space $\mathrm{L}^2(\mathbb{X},\nu)^{\circ n}$ is endowed by the scalar product

$$\langle f_n, g_n\rangle_{\mathrm{L}^2(\mathbb{X},\nu)^{\circ n}} = n!\int_{\mathbb{X}^{n,<}} f_n(\mathbf{t}_n,\mathbf{k}_n)\,g_n(\mathbf{t}_n,\mathbf{k}_n)\,\mathrm{d}\nu^{\otimes n}(\mathbf{t}_n,\mathbf{k}_n),$$

where the tensor measure $\nu^{\otimes n}$ is defined on $\mathbb{X}^{n,\neq}$ by $\nu^{\otimes n} = \bigotimes_{i=1}^n \nu$.

The multiple stochastic integral can be defined on $\mathcal{C}_c(\mathbb{X}^n,\mathbf{R})$, the set of continuous functions with compact support on $\mathbb{X}^n$ and extended by isometry to $\mathrm{L}^2(\mathbb{X},\nu)^{\circ n}$.

**Proposition 2.6.** *The $\mathcal{R}$-stochastic integral of order $n$ is the application defined on $\mathcal{C}_c(\mathbb{X}^n,\mathbf{R})$ by*

$$\mathrm{J}_n(f_n\,;\mathcal{R}) = n\sum_{(t,k)\in\mathbb{X}} \mathrm{J}_{n-1}(f_n(\star,(t,k)))\,\Delta\mathrm{R}_{(t,k)}, \tag{2.7}$$

*where "$\star$" denotes the first $n-1$ variables of $f_n((t_1,k_1),\dots,(t_n,k_n))$. It can equivalently be written as*

$$\mathrm{J}_n(f_n\,;\mathcal{R}) = n!\sum_{(\mathbf{t}_n,\mathbf{k}_n)\in\mathbb{X}^n} f_n(\mathbf{t}_n,\mathbf{k}_n)\prod_{i=1}^n \Delta\mathrm{R}_{(t_i,k_i)}. \tag{2.8}$$

*Besides, it satisfies the isometry formula: for any $f_n\in\mathrm{L}^2(\mathbb{X},\nu)^{\circ n}$, $g_m\in\mathrm{L}^2(\mathbb{X},\nu)^{\circ m}$*

$$\mathbf{E}\left[\mathrm{J}_n(f_n\,;\,\mathcal{R})\mathrm{J}_m(g_m\,;\,\mathcal{R})\right] = \mathbf{1}_{\{n\}}(m)n!\langle f_n, g_n\rangle_{\mathrm{L}^2(\mathbb{X},\tilde{\nu})^{\circ n}}, \tag{2.9}$$

*so that its domain can be extended to $\mathrm{L}^2(\mathbb{X},\mathcal{X},\tilde{\nu})^{\circ n}\simeq\mathrm{L}^2(\mathbb{X},\mathcal{X},\nu)^{\circ n}$.*

Up to now, if no need to specify, $\mathrm{L}^2(\mathbb{X})^{\circ n}$ could indifferently designate $\mathrm{L}^2(\mathbb{X},\mathcal{X},\nu)^{\circ n}$ or $\mathrm{L}^2(\mathbb{X},\mathcal{X},\tilde{\nu})^{\circ n}$. This subsection ends up with two Lemmas that will be useful to state the chaotic expansion theorem.

**Lemma 2.7.** *For any $g \in \mathrm{L}^2(\mathbb{X})$ and $f_n \in \mathrm{L}^2(\mathbb{X})^{\circ n}$,*

$$\mathrm{J}_{n+1}(g \circ f_n\,;\mathcal{R}) = n \sum_{(t,k)\in\mathbb{X}} \mathrm{J}_n\big(f_n(\star,(t,k)) \circ g(\cdot)\mathbf{1}_{\{1:t-1\}^n}(\star,\cdot)\,;\mathcal{R}\big)\Delta\mathrm{R}_{(t,k)} + \sum_{(t,k)\in\mathbb{X}} g(t,k)\mathrm{J}_n(f_n\mathbf{1}_{\{1:t-1\}^n}\,;\mathcal{R})\Delta\mathrm{R}_{(t,k)},$$

*where $\circ$ designates the symmetric tensor product and satisfies for $(\mathbf{t}_n,\mathbf{k}_n) \in \mathbb{X}^n$,*

$$g \circ f_n(\mathbf{t}_{n+1},\mathbf{k}_{n+1}) = \frac{1}{n+1}\sum_{i=1}^{n+1} g(t_i,k_i) f_n^{\neg i}(\mathbf{t}_n,\mathbf{k}_n),$$

*with for $i \in \{1:n\}$, $f_n^{\neg i}(\mathbf{t}_n,\mathbf{k}_n) = f_n\big((t_1,k_1),\cdots,(t_{i-1},k_{i-1}),(t_{i+1},k_{i+1}),(t_n,k_n)\big)$.*

**Lemma 2.8.** *For any $(t,n) \in \mathbf{N}^2$, $f_n \in \mathrm{L}^2(\mathbb{X})^{\circ n}$,*

$$\mathbf{E}\left[\mathrm{J}_n(f_n\,;\mathcal{R})\,|\,\mathcal{F}_t\right] = \mathrm{J}_n(f_n\mathbf{1}_{\{1:t\}}\,;\mathcal{R}).$$

## 2.3 Chaotic decomposition

This subsection is devoted to the statement of a chaotic decomposition for any square-integrable *marked binomial functional*, that are random variables of the form

$$\mathrm{F} = f_0\mathbf{1}_{\{\eta(\mathbb{X})=0\}} + \sum_{n\in\mathbf{N}}\sum_{(\mathbf{t}_n,\mathbf{k}_n)\in\mathbb{X}^n} \mathbf{1}_{\{\eta(\mathbb{X})=n\}} f_n(\mathbf{t}_n,\mathbf{k}_n)\prod_{i=1}^{n}\mathbf{1}_{\{(t_i,k_i)\in\eta\}}, \tag{2.10}$$

where any function $f_n$ is an element of $\mathrm{L}^1(\nu)^{\circ n}$, that is the subspace of $\mathrm{L}^1(\nu)^{\otimes n} := \mathrm{L}^1(\mathbb{X},\mathcal{X},\nu)^{\otimes n} = \mathrm{L}^1(\mathbb{X},\mathcal{X},\nu)^n$ composed of the functions symmetric in their $n$ variables. We introduce the space of cylindrical functions, which is dense in $\mathrm{L}^2(\mathbf{P})$.

**Definition 2.9.** A functional F is *cylindrical* if there exists $T \in \mathbf{N}$ such that

$$\mathrm{F} = f_0\mathbf{1}_{\{\eta(\mathbb{X})=0\}} + \sum_{n\in\mathbf{N}}\sum_{(\mathbf{t}_n,\mathbf{k}_n)\in\mathbb{X}_T^n} \mathbf{1}_{\{\eta(\mathbb{X})=n\}} f_n(\mathbf{t}_n,\mathbf{k}_n)\prod_{i=1}^{n}\mathbf{1}_{\{(t_i,k_i)\in\eta\}}, \tag{2.11}$$

where $\mathbb{X}_T = \{(t,k) \in \mathbb{X} \,:\, t \leqslant T\}$.

Within Assumption 2.2, let $\mathcal{H}_0 := \mathbf{R}$ and for any $n \in \mathbf{N}$, $\mathcal{H}_n$ be the subspace of $\mathrm{L}^2(\mathbf{P})$ made of integrals of order $n \geqslant 1$:

$$\mathcal{H}_n := \left\{\mathrm{J}_n(f_n)\,;\; f_n \in \mathrm{L}^2(\mathbb{X})^{\circ n}\right\},$$

where $\mathrm{J}_n(f_n) := \mathrm{J}_n(f_n\,;\mathcal{R})$, and called *chaos of order $n$*. In what follows $\mathcal{L}^0(\mathbf{P},\mathcal{F}_t)$ denotes the set of $\mathcal{F}_t$-measurable random variables.

**Lemma 2.10.** *For any $t \in \mathbf{N}$,*

$$\mathcal{L}^0(\mathbf{P},\mathcal{F}_t) = (\mathcal{H}_0 \oplus \cdots \oplus \mathcal{H}_t)\bigcap \mathcal{L}^0(\mathbf{P},\mathcal{F}_t). \tag{2.12}$$

As a direct consequence, any random variable $\mathrm{F} \in \mathcal{L}^0(\mathbf{P}, \mathcal{F}_t)$ can be expressed as

$$\mathrm{F} = \mathbf{E}[\mathrm{F}] + \sum_{n=1}^{t} \mathrm{J}_n\big(f_n \mathbf{1}_{\{1:t\}^n}\big).$$

This also means that the space of cylindrical functions coincides with the linear space spanned by multiple stochastic integrals i.e.

$$\mathcal{S} = \mathrm{Span}\Big\{ \bigcup_{n\in\mathbf{Z}_+} \mathcal{H}_n \Big\}.$$

The completion of $\mathcal{S}$ in $\mathrm{L}^2(\mathbf{P})$ is denoted by the sum $\bigoplus_{n\in\mathbf{Z}_+} \mathcal{H}_n$. We can state the main theorem of this section and provide a chaotic decomposition for any square-integrable random variable.

**Theorem 2.11.** *The space of square-integrable marked binomial functionals is provided with a chaotic decomposition*

$$\mathrm{L}^2(\mathbf{P}) = \bigoplus_{n\in\mathbf{Z}_+} \mathcal{H}_n. \tag{2.13}$$

*In other terms, any random variable* $\mathrm{F} \in \mathrm{L}^2(\mathbf{P})$ *can be expanded in a unique way as*

$$\mathrm{F} = \mathbf{E}[\mathrm{F}] + \sum_{n\in\mathbf{N}} \mathrm{J}_n(f_n). \tag{2.14}$$

*Proof.* The proof follows closely that of Proposition 1.5.3 in Privault [56] by combining Lemma 2.10 and the density of $\mathcal{S}$ in $\mathrm{L}^2(\mathbf{P})$. □

**Corollary 2.12.** *For any* $\mathrm{F}, \mathrm{G} \in \mathrm{L}^2(\mathbf{P})$,

$$\mathrm{cov}(\mathrm{F}, \mathrm{G}) = \sum_{n\in\mathbf{N}} n! \langle f_n, g_n \rangle_{\mathrm{L}^2(\mathbb{X},\tilde{\nu})^{\otimes n}}.$$

*Proof.* Immediate using (2.14) together with Proposition 2.4. □

**Remark 2.13.** The chaotic decomposition for marked binomial processes is not generated - as in the framework of normal martingales (including Brownian motion, Poisson and Rademacher processes) - from the increments of the compensated underlying process $\overline{\mathrm{Y}}$ (2.4) itself but in terms of stochastic integrals with respect to an auxiliary process R satisfying both a martingale and an orthogonality properties (Assumption 2.2). This can be explained by the absence of normal martingales associated to the compound binomial process $\overline{\mathrm{Y}}$. Indeed, by transposing the remark p. 95 in Privault [56] into our framework, the quadratic variation of the compensated compound $\mathcal{F}$-martingale $\overline{\mathrm{Y}}$ satisfies

$$[\overline{\mathrm{Y}}, \overline{\mathrm{Y}}]_t = \frac{1}{\sqrt{\lambda \mathrm{var}[\mathrm{V}_1]}} \sum_{s=1}^{\mathrm{N}_t} |\mathrm{V}_s|^2 = \frac{1}{\sqrt{\lambda \mathrm{var}[\mathrm{V}_1]}} \sum_{s=1}^{t} |\mathrm{V}_{\mathrm{N}_s}| \, \Delta\overline{\mathrm{Y}}_s + \frac{\mathbf{E}[\mathrm{V}_1]}{\sqrt{\mathrm{var}[\mathrm{V}_1]}} \sum_{s=1}^{t} |\mathrm{V}_{\mathrm{N}_s}|,$$

does not allow to find a square-integrable $\mathcal{F}$-adapted process $(\phi_t)_{t\in\mathbf{R}_+}$ solution of the *structure equation*

$$[\overline{\mathrm{Y}}, \overline{\mathrm{Y}}]_t = t + \sum_{s=1}^{t} \phi_s \, \Delta\overline{\mathrm{Y}}_s,$$

when V is not deterministic. This structural reason explains the lack of usual chaotic decomposition with respect to the increments of the compensated initial process.

Despite previous remark, we can nevertheless provide a *pseudo*-chaotic (not orthogonal) decomposition related to the process Y. In order to do that, we introduce the process $\mathrm{Z} = (\mathrm{Z}_{(t,k)}\,;\,(t,k) \in \mathbb{X})$ which increments are defined by the family $\mathcal{Z} = \{\Delta\mathrm{Z}_{(t,k)}\,;\,(t,k) \in \mathbb{X}\}$ with

$$\Delta\mathrm{Z}_{(t,k)} = \mathbf{1}_{\{(\Delta\mathrm{N}_t,\mathrm{V}_{\mathrm{N}_t})=(t,k)\}} - \lambda\mathbf{Q}(\{k\}) = \mathbf{1}_{\{(\Delta\mathrm{N}_t,\mathrm{W}_t)=(t,k)\}} - \lambda\mathbf{Q}(\{k\})\ ;\ (t,k) \in \mathbb{X}. \tag{2.15}$$

The definition of $\mathcal{Z}$ is quite natural since the process $\overline{\mathrm{Y}} = (\overline{\mathrm{Y}}_t)_{t\in\mathbf{N}}$ can be equivalently written

$$\overline{\mathrm{Y}}_t = \sum_{s\leqslant t}\sum_{k\in\mathrm{E}} k\Delta\mathrm{Z}_{(s,k)}. \tag{2.16}$$

For any $T \in \mathbf{N}$, define $\mathcal{Z}_T = \{\Delta\mathrm{Z}_{(t,k)}\,;\,(t,k) \in \mathbb{X}_T\}$. This family is not orthogonal; however, the finite dimension of the related spanned space, being equal to

$$1 + \sum_{s=1}^{T} |\mathrm{E}|^s \times \binom{T}{s} = (|\mathrm{E}| + 1)^T =: \overline{\mathfrak{m}},$$

enables to derive from $\mathcal{Z}_T$ an orthogonal family, $\mathcal{R}_T = \{\Delta\mathrm{R}_{(t,k)}\,;\,(t,k) \in \mathbb{X}_T\}$. Assume that $\mathrm{E} = \{k^1, \cdots, k^{\overline{\mathfrak{m}}})$; then, the Gram-Schmidt process provides

$$\Delta\mathrm{R}_0 = 1, \quad \Delta\mathrm{R}_{(t,k^1)} = \Delta\mathrm{Z}_{(1,k^1)} \quad \text{and} \qquad \Delta\mathrm{R}_{(t,k^n)} = \Delta\mathrm{Z}_{(t,k^n)} - \sum_{j=1}^{n-1} \frac{\mathbf{E}\left[\Delta\mathrm{Z}_{(1,k^n)}\Delta\mathrm{R}_{(1,k^j)}\right]}{\mathbf{E}\left[(\Delta\mathrm{R}_{(1,k^j)})^2\right]}\Delta\mathrm{R}_{(t,k^j)}, \tag{2.17}$$

for $n \in \{1 : \overline{\mathfrak{m}}\}$, by noting that the random variables $\Delta\mathrm{R}_{(t,k)}$ (respectively $\Delta\mathrm{Z}_{(t,k)}$) and $\Delta\mathrm{R}_{(1,k)}$ (respectively $\Delta\mathrm{Z}_{(1,k)}$) are identically distributed and that for any $s \in \{1 : t-1\}$, $\mathbf{E}\left[\Delta\mathrm{R}_{(s,k)}\,\Delta\mathrm{Z}_{(t,1)}\right] = \mathbf{E}\left[\Delta\mathrm{R}_{(s,k)}\mathbf{E}\left[\Delta\mathrm{Z}_{(t,1)}|\mathcal{F}_s\right]\right] = 0$. In fact, for any $t \in \{1 : T\}$, $(\Delta\mathrm{Z}_{(t,k)},\ k \in \mathrm{E})$ is the image of $(\Delta\mathrm{R}_{(t,k)},\ k \in \mathrm{E})$ by the linear transformation of associated to the $\overline{\mathfrak{m}} \times \overline{\mathfrak{m}}$ triangular matrix

$$\mathfrak{M} = (\mathfrak{m}_{ij})_{i,j\in\{1:\overline{\mathfrak{m}}\}} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ \gamma_{21} & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{n1} & \gamma_{n2} & \cdots & 1 \end{pmatrix}, \tag{2.18}$$

where $\gamma_{ij} := \mathbf{E}\left[\Delta\mathrm{Z}_{(1,k^i)}\Delta\mathrm{R}_{(1,k^j)}\right] / \mathbf{E}\left[(\Delta\mathrm{R}_{(1,k^j)})^2\right]$. As the matrix $\mathfrak{M}$ is invertible, for any $t \in \{1 : T\}$, $(\Delta\mathrm{R}_{(t,k)},\ k \in \mathrm{E})$ is obtained through the product of $\mathfrak{M}^{-1}$ by the vector $(\Delta\mathrm{Z}_{(t,k)},\ k \in \mathrm{E})$. Moreover, since the linear transformation it stands for is then bijective, the family $\mathcal{R}$ can be constructed in a similar fashion when E is countable not finite. Thus, the process R which increments are defined by the family $\mathcal{R}$ satisfies Assumption 2.2.

**Remark 2.14.** It seems to be possible to construct such a family $\mathcal{R}$ even if E is not countable (take for instance $\mathrm{E} = \mathbf{R}$), by drawing inspiration from the design of the *orthogonal power jump process* for Lévy processes, in Di Nunno, Oksendal and Proske [52]. Transposing it into our framework, that would give: define for any $n \in \mathbf{N}$,

$$\Delta\mathrm{Z}_t^{(n)} = \mathrm{X}_t^{(n)} - \mathbf{E}\big[\mathrm{X}_t^{(n)}\big] := \sum_{s\in\{1:t\}} (\Delta\mathrm{Y}_s)^n - \mathbf{E}\Bigg[\sum_{s\in\{1:t\}} (\Delta\mathrm{Y}_s)^n\Bigg],$$

and the family $\mathcal{R}$ by $\Delta \mathrm{R}_0 = 1$, and

$$\Delta \mathrm{R}_t^{(n)} = \mathrm{X}_t^{(n)} + \sum_{j=1}^{n-1} \gamma_{nj} \mathrm{X}_t^{(j)},$$

where the $\gamma_{nj}$ are real numbers such that the processes of the collection $\{(\Delta \mathrm{R}_t^{(n)})_{t\in\mathbf{N}},\ n \in \mathbf{N}\}$ are *strongly orthogonal martingales*, i.e. for any $t \in \mathbf{N}$, the product $\Delta \mathrm{R}^{(n)} \Delta \mathrm{R}^{(m)}$ is a uniformly integrable martingale for all $(n, m) \in \mathbf{N}^2$, $m \neq n$.

**Remark 2.15.** Let the $\mathcal{Z}$-stochastic integral of order $n \in \mathbf{N}$ be the application on $\mathrm{L}^2(\mathbb{X})^{\circ n}$ such that for any $f_n \in \mathrm{L}^2(\mathbb{X})^{\circ n}$,

$$\mathrm{J}_n(f_n\,; \mathcal{Z}) := \sum_{(\mathbf{t}_n,\mathbf{k}_n)\in\mathbb{X}^n} f_n(\mathbf{t}_n, \mathbf{k}_n) \prod_{i=1}^{n} \Delta \mathrm{Z}_{(t_i,k_i)}.$$

Considering the application $\mathbf{1}^{<}_{(\mathbf{t}_n,\mathbf{k}_n)} : (\mathbf{s}_n, \mathbf{l}_n) \in \mathbb{X}^{n,<} \mapsto \mathbf{1}_{(\mathbf{t}_n,\mathbf{k}_n)}(\mathbf{s}_n, \mathbf{l}_n)$, we retrieve the remarkable and usual identity

$$\mathrm{J}_n\big(\mathbf{1}^{<}_{(\mathbf{t}_n,\mathbf{k}_n)}\,; \mathcal{Z}\big) = \prod_{i=1}^{n} \Delta \mathrm{Z}_{(t_i,k_i)}\ ;\ n \in \mathbf{N}.$$

This is of key importance; it basically means that we can reconstruct the signal Y by means of the stochastic integral of elementary functions defined on $\mathbb{X}^n$. In particular for $n = 1$, this gives $\mathrm{J}_1(\mathbf{1}_{(t,k)}\,; \mathcal{Z}) = \Delta \mathrm{Z}_{(t,k)}$, that appears as a reminiscence of the Wiener integral.

We can thus state the following result.

**Proposition 2.16.** *Any random variable* $\mathrm{F} \in \mathrm{L}^2(\mathbf{P})$ *can be expressed as*

$$\mathrm{F} = \mathbf{E}\,[\mathrm{F}] + \sum_{n\in\mathbf{N}} \mathrm{J}_n(g_n, ; \mathcal{Z}). \tag{2.19}$$

*In particular if* $|\mathrm{E}| = \overline{\mathfrak{m}}$ *such that* $\mathrm{E} = \{k^1, \ldots, k^{\overline{\mathfrak{m}}}\}$*, any function* $g_n$ *is explicitely given by*

$$g_n(\mathbf{t}_n, \mathbf{k}_n) = \sum_{i_1=p_1}^{\overline{\mathfrak{m}}} \cdots \sum_{i_n=p_n}^{\overline{\mathfrak{m}}} \Big( \prod_{j=1}^{n} \mathfrak{m}^{-1}_{k^{i_j} k^{p_j}} \Big) f_n((t_1, k_1^{i_1}), \ldots, (t_n, k_n^{i_n})),$$

*where for any* $j \in \{1, \ldots, n\}$*,* $p_j$ *denotes the element of* $\{1, \ldots, m\}$ *such that* $k_j = k^{p_j} \in \mathrm{E}$*, and for notation purposes,* $\mathfrak{m}^{-1}_{k^i,k^j}$ *designate the* $(i, j)$*-th entry of matrix* $\mathfrak{M}^{-1}$*, the inverse of matrix* $\mathfrak{M}$ *defined by* (2.18)*.*

## 2.4 Doléans exponentials

Define for any $h \in \mathrm{L}^2(\mathbb{X})$ the exponential vector by

$$\xi(h) = \mathbf{E}\,[\xi(h)] + \sum_{n\in\mathbf{N}} \frac{1}{n!} \mathrm{J}_n(h^{\otimes n}). \tag{2.20}$$

The family $(\xi_t(h))_{t\in\mathbf{N}}$ defined by $\xi_t(h) = \xi(h\mathbf{1}_{\{1:t\}})$ can be viewed as a discrete Doléans exponential solution of the equation in differences

$$\xi_t(h) - \xi_{t-1}(h) = \xi_{t-1}(h)\sum_{k\in\mathrm{E}} g(t,k)\Delta\mathrm{Z}_{(t,k)},\ t\in\mathbf{N},$$

where $g\in\mathrm{L}^2(\mathbb{X})$ is given in the following theorem.

**Proposition 2.17.** *For any $h\in\mathrm{L}^2(\mathbb{X})$, the discrete Doléans exponential defined by* (2.20) *can be written as*

$$\xi(h) = \mathbf{E}\left[\xi(h)\right]\prod_{t\in\mathbf{N}}\Big(1+\sum_{k\in\mathrm{E}} g(t,k)\big(\mathbf{1}_{(t,k)}-\lambda\mathbf{Q}(\{k\})\big)\Big), \tag{2.21}$$

*where $g$ is the element of* $\mathrm{L}^2(\mathbb{X})$ *such that* $\mathrm{J}_1(g\,;\mathcal{Z}) = \mathrm{J}_1(h)$.

# 3 Stochastic analysis for marked binomial processes, part II

The section is organised as follows; the first subsection is dedicated to the development of a $\mathrm{L}^1$-theory for binomial marked processes which starting point is a Mecke-type formula. In the just following part, are provided some elements of Malliavin calculus whereas in the third subsection, the tools of $\mathrm{L}^1$ and $\mathrm{L}^2$ theories are gathered to formalize a unified Markov-Malliavin structure.

## 3.1 $\mathrm{L}^1$-theory: the Mecke and Mehler's formulas

### 3.1.1 The Mecke formula and difference operators on $\mathrm{L}^1$

The following Lemma is the analogue of the Mecke formula for marked binomial processes.

**Lemma 3.1.** *Let $\eta$ be a marked binomial process on $\mathbb{X}$ with intensity measure $\nu$. Then for any real-valued, non-negative, $\mathbb{X}\times\mathfrak{N}_{\mathbb{X}}$-measurable function $\mathfrak{u}$,*

$$\mathbf{E}\left[\sum_{(t,k)\in\eta}\mathfrak{u}(\eta,(t,k))\right] = \mathbf{E}\left[\int_{\mathbb{X}}\mathfrak{u}\big(\pi_t(\eta)+\delta_{(t,k)},(t,k)\big)\mathrm{d}\nu(t,k)\right]. \tag{3.1}$$

*where the application $\pi_t : \mathfrak{N}_{\mathbb{X}}\to\mathfrak{N}_{\mathbb{X}}$ is the restriction of $\eta$ to $\mathcal{G}_t := \sigma\big\{\sum_{(s,k)}\eta(s,k), s\neq t, k\in\mathrm{E}\big\}$ i.e.*

$$\pi_t(\eta) = \sum_{s\neq t}\sum_{k\in\mathrm{E}}\eta(s,k). \tag{3.2}$$

**Remark 3.2.** Clearly, the formula (3.1) continues to hold provided the process $u$ of representative $\mathfrak{u}$ belongs to $\mathrm{L}^1(\mathbf{P}\otimes\nu)$.

The applications defined on $\mathfrak{N}_{\mathbb{X}}\times\mathbb{X}$, and expressed for any $\big(\eta,(t,k)\big)\in\mathfrak{N}_{\mathbb{X}}\times\mathbb{X}$ by

$$\eta\mapsto\pi_t(\eta)+\delta_{(t,k)}\quad\text{and}\quad\eta\mapsto\pi_t(\eta), \tag{3.3}$$

can be interpreted as the applications acting on $\eta$ respectively by forcing a jump of height $k$ at time $t$ or forbidding any jump at time $t$. As a reminiscence of Poisson space theory, define $\mathrm{D}^+$ the *add-one cost operator* for any $\mathrm{F} \in \mathcal{L}^0(\Omega)$ by

$$\mathrm{D}^+_{(t,k)}\mathrm{F} := \mathfrak{f}(\pi_t(\eta) + \delta_{(t,k)}) - \mathfrak{f}(\pi_t(\eta)). \tag{3.4}$$

The difference operator $\mathrm{D}^+$ measures the effect of adding a point $(t,k) \in \mathbb{X}$ to $\eta$ compared to the process truncated at time $t$. The product formula can be easily deduced from this expression of D and is strongly reminiscent to that existing in the Poisson setting (see for instance Privault [56], Proposition 6.4.8). For $\mathrm{F}, \mathrm{G} \in \mathrm{L}^1(\mathbf{P})$ of respective representatives $\mathfrak{f}$ and $\mathfrak{g}$,

$$\mathrm{D}^+_{(t,k)}(\mathrm{FG}) = \mathfrak{f}(\pi_t(\eta))(\mathrm{D}^+_{(t,k)}\mathrm{G}) + \mathfrak{g}(\pi_t(\eta))(\mathrm{D}^+_{(t,k)}\mathrm{F}) + (\mathrm{D}^+_{(t,k)}\mathrm{F})(\mathrm{D}^+_{(t,k)}\mathrm{G}). \tag{3.5}$$

**Remark 3.3.** By definition, given $k \in \mathrm{E}$, for any $t \in \mathbf{N}$ the random variables $\pi_t(\eta) + \delta_{(t,k)}$ and $\pi_t(\eta)$ are $\mathcal{G}_t$-measurable; so does $\mathrm{D}^+_{(t,k)}\mathrm{F}$.

In a similar way the operator $\mathrm{D}^-$ is defined for any $\mathrm{F} \in \mathcal{L}^0(\Omega)$ by

$$\mathrm{D}^-_{(t,k)}\mathrm{F} := \mathfrak{f}(\eta) - \mathfrak{f}(\eta - \delta_{(t,k)}), \tag{3.6}$$

if $(t,k) \in \eta$, and is equal to zero otherwise. The operator $\mathrm{D}^-$ may be interpreted as a *remove-one cost operator*: if the point $(t,k)$ was charged by $\eta$, this is removed by the action of $\mathrm{D}^-_{(t,k)}$. The operator $\mathrm{D}^-$ satisfies the product formula: for $\mathrm{F}, \mathrm{G} \in \mathcal{L}^0(\Omega)$,

$$\mathrm{D}^-_{(t,k)}(\mathrm{FG}) = \mathrm{F}(\mathrm{D}^-_{(t,k)}\mathrm{G}) + \mathrm{G}(\mathrm{D}^-_{(t,k)}\mathrm{F}) - (\mathrm{D}^-_{(t,k)}\mathrm{F})(\mathrm{D}^-_{(t,k)}\mathrm{G}). \tag{3.7}$$

Define on $\mathrm{L}^1(\mathbf{P} \otimes \nu)$ the operator $\tilde{\delta}$ such that for any process $u \in \mathrm{L}^1(\mathbf{P} \otimes \nu)$ of representative $\mathfrak{u}$,

$$\tilde{\delta}(u) := \sum_{(t,k)\in\eta} \mathfrak{u}(\eta,(t,k)) - \int_{\mathbb{X}} \mathfrak{u}(\eta,(t,k))\mathrm{d}\nu(t,k). \tag{3.8}$$

As $\pi_t(\eta) + \delta_{(t,k)} = \eta$ if $(t,k) \in \eta$, we can additionally introduce the operator $\widetilde{\mathrm{L}}$ on $\mathcal{L}^0(\Omega)$ such that

$$\begin{aligned} \widetilde{\mathrm{L}}\mathrm{F} := -\tilde{\delta}(\mathrm{D}^+\mathrm{F}) &= -\sum_{(t,k)\in\eta} \left[\mathfrak{f}(\pi_t(\eta) + \delta_{(t,k)}) - \mathfrak{f}(\eta)\right] + \int_{\mathbb{X}} \left[\mathfrak{f}(\pi_t(\eta) + \delta_{(t,k)}) - \mathfrak{f}(\pi_t(\eta))\right]\mathrm{d}\nu(t,k) \\ &= -\sum_{(t,k)\in\eta} \left[\mathfrak{f}(\eta) - \mathfrak{f}(\eta - \delta_{(t,k)})\right] + \int_{\mathbb{X}} \left[\mathrm{D}^+_{(t,k)}\mathrm{F}\right] \mathrm{d}\nu(t,k) \\ &= \sum_{(t,k)\in\eta} \left[\mathrm{D}^-_{(t,k)}\mathrm{F}\right] + \int_{\mathbb{X}} \left[\mathrm{D}^+_{(t,k)}\mathrm{F}\right] \mathrm{d}\nu(t,k), \end{aligned} \tag{3.9}$$

for any $\mathrm{F} \in \mathcal{L}^0(\Omega)$. The Mecke equation (3.1) ensures that this definition does not depend $\mathbf{P}$-a.s. on the choice of the representative. We get the following "almost"-$\mathrm{L}^1$-integration by parts formula.

**Proposition 3.4.** *For any predictable process* $u \in \mathcal{L}^0(\Omega \times \mathbf{N})$ *and* $\mathrm{F} \in \mathcal{L}^0(\Omega)$*,*

$$\mathbf{E}\left[\int_{\mathbb{X}} \mathrm{D}^+\mathrm{F}\, \mathfrak{u}(\eta,(t,k))\, \mathrm{d}\nu(t,k)\right] = \mathbf{E}[\mathrm{F}\widetilde{\delta}(u)] + \mathbf{E}\left[\int_{\mathbb{X}} \overline{\mathrm{D}}_t\mathrm{F}\, \mathfrak{u}(\eta,(t,k))\, \mathrm{d}\nu(t,k)\right].$$

*where the operator* $\overline{\mathrm{D}}$ *is defined on* $\mathcal{L}^0(\Omega)$ *by*

$$\overline{\mathrm{D}}_t(\mathrm{F}) = \mathfrak{f}(\eta) - \mathfrak{f}(\pi_t(\eta))\ ;\ t \in \mathbf{N}.$$

**Remark 3.5.** This latter can be rewritten as

$$\mathbf{E}\big[\langle \widetilde{\mathrm{D}}\mathrm{F}, u\rangle_{\mathrm{L}^2(\mathbb{X},\nu)}\big] = \mathbf{E}\big[\mathrm{F}\widetilde{\delta}(u)\big], \tag{3.10}$$

where $\widetilde{\mathrm{D}}_{(t,k)}\mathrm{F} = \mathrm{D}^+_{(t,k)}\mathrm{F} - \overline{\mathrm{D}}_t\mathrm{F} = \mathfrak{f}(\pi_t(\eta) + \delta_{(t,k)}) - \mathfrak{f}(\eta)$. The operator $\widetilde{\mathrm{D}}$ is the exact discrete analogue of the usual gradient on Poisson space. In that latter case, provided the intensity measure of the Poisson point process is diffuse, $\mathrm{D}^+$ and $\widetilde{\mathrm{D}}$ are equal $\mathbf{P}\otimes\nu$ almost surely. That does not hold here, and we justify our choice to define the add-one-cost operator by $\mathrm{D}^+$ and not via $\widetilde{\mathrm{D}}$ in the perspective to combine $\mathrm{L}^1$ and $\mathrm{L}^2$ later on through (3.20). This remark is crucial to understand why Gamma calculus can not perform within our framework. Indeed, let us introduce the operator $\widetilde{\Gamma}$ defined for any random variables $\mathrm{F}, \mathrm{G} \in \mathcal{L}^0(\Omega)$ such that $(\mathrm{D}^+\mathrm{F})(\mathrm{D}^+\mathrm{G}) \in \mathrm{L}^1(\mathbf{P}\otimes\nu)$, by $\widetilde{\Gamma}(\mathrm{F},\mathrm{G}) = 1/2\big[\widetilde{\mathrm{L}}(\mathrm{FG}) - \mathrm{F}(\widetilde{\mathrm{L}}\mathrm{G}) - \mathrm{G}(\widetilde{\mathrm{L}}\mathrm{F})\big]$. By combining the product rules (3.5) and (3.7), we obtain

$$\begin{aligned}\widetilde{\Gamma}(\mathrm{F},\mathrm{G}) = \frac{1}{2}\bigg[&\int_{\mathbb{X}} (\mathrm{D}^+_{(t,k)}\mathrm{F})(\mathrm{D}^+_{(t,k)}\mathrm{G})\, \mathrm{d}\nu(t,k) + \int_{\mathbb{X}} (\mathrm{D}^-_{(t,k)}\mathrm{F})(\mathrm{D}^-_{(t,k)}\mathrm{G})\, \mathrm{d}\eta(t,k) \\ &- \int_{\mathbb{X}} (\mathrm{D}^+_{(t,k)}\mathrm{F})(\overline{\mathrm{D}}_t\mathrm{G})\, \mathrm{d}\nu(t,k) - \int_{\mathbb{X}} (\mathrm{D}^+_{(t,k)}\mathrm{G})(\overline{\mathrm{D}}_t\mathrm{F})\, \mathrm{d}\nu(t,k)\bigg],\end{aligned} \tag{3.11}$$

whereas, as a consequence of the Mecke formula we can prove that for any $\mathrm{F}, \mathrm{G} \in \mathcal{L}^0(\Omega)$ of respective representatives $\mathfrak{f}$ and $\mathfrak{g}$ such that $\mathrm{F}\mathfrak{g}(\pi_t(\eta)+\delta_\cdot)$, $\mathfrak{f}(\pi_t(\eta)+\delta_\cdot)\mathrm{G}$, $\mathfrak{f}(\pi_t(\eta)+\delta_\cdot)\mathfrak{g}(\pi_t(\eta)+\delta_\cdot) \in \mathrm{L}^1(\mathbf{P}\otimes\nu)$, we have

$$\begin{aligned}-\mathbf{E}\big[\widetilde{\Gamma}(\mathrm{F},\mathrm{G})\big] &= \frac{1}{2}\Big[\mathbf{E}[\mathrm{F}(\widetilde{\mathrm{L}}\mathrm{G})] + \mathbf{E}[\mathrm{G}(\widetilde{\mathrm{L}}\mathrm{F})]\Big] \\ &= \frac{1}{2}\mathbf{E}\left[\left(\int_{\mathbb{X}} (\mathrm{D}^+_{(t,k)}\mathrm{G})(\overline{\mathrm{D}}_t\mathrm{F})\mathrm{d}\nu(t,k) + \int_{\mathbb{X}} (\mathrm{D}^+_{(t,k)}\mathrm{F})(\overline{\mathrm{D}}_t\mathrm{G})\mathrm{d}\nu(t,k)\right)\right],\end{aligned}$$

from which it is not possible to draw an $\mathrm{L}^1$-integration by parts formula since $\mathbf{E}[\mathrm{F}(\widetilde{\mathrm{L}}\mathrm{G})] \neq \mathbf{E}[\mathrm{G}(\widetilde{\mathrm{L}}\mathrm{F})]$. As a result, the possibility to combine $\mathrm{L}^1$ and $\mathrm{L}^2$ theories in regards of the carré du champ operator $\Gamma$ seems compromised; their connection can at best come at the level of the operator L.

## 3.2 $\mathrm{L}^2$-theory: Malliavin operators

From the chaotic decomposition that equips the space $\mathrm{L}^2(\mathbf{P})$, we define the Malliavin operators, *gradient*, *divergence*, *number operator*, and the *Ornstein-Uhlenbeck semi-goup*.

### 3.2.1 Gradient

As one way to develop it, we introduce the Malliavin derivative as the *annihilation operator* acting on the space $\mathrm{L}^2(\mathbf{P})$ seen in terms of its chaotic expansion (2.13).

**Definition 3.6.** Let $\mathbf{D}_0$ be the set of random variables $\mathrm{F} \in \mathrm{L}^2(\mathbf{P})$ whose decomposition (2.14) satisfies

$$\sum_{n=1}^{\infty} nn! \|f_n\|^2_{\mathrm{L}^2(\mathbb{X})^{\otimes n}} < \infty. \tag{3.12}$$

Let the linear, unbounded, closable operator $\mathrm{D} : \mathbf{D}_0 \to \mathrm{L}^2(\mathbf{P} \otimes \nu)$ be defined for any element $\mathrm{J}_n(f_n)$ of $\mathcal{H}_n$ by

$$\mathrm{D}_{(t,k)} \mathrm{J}_n(f_n) = n\, \mathrm{J}_{n-1}\big(f_n(\star, (t,k)) \mathbf{1}_{\{1:n-1\}^{n,<}}\big). \tag{3.13}$$

### 3.2.2 Divergence

Let $\mathcal{U}$ be the space

$$\mathcal{U} = \Bigg\{ \sum_{n \in \{0:T\}} \mathrm{J}_n(f_{n+1}(\star, \cdot))\, ; \, f_{n+1} \in \mathrm{L}^2(\mathbb{X})^{\circ n} \otimes \mathrm{L}^2(\mathbb{X}),\, n \in \{0:T\},\, T \in \mathbf{N} \Bigg\}. \tag{3.14}$$

The operator divergence is introduced as the *creation operator* acting on $\mathrm{L}^2(\mathbf{P})$, that can be, thanks to Theorem 2.11, understood as a Fock space.

**Definition 3.7.** Let the linear, unbounded, closable operator $\delta : \operatorname{dom} \delta \to \mathrm{L}^2(\mathbf{P})$ whose domain $\operatorname{dom} \delta$ (that will be described later) contains the set of processes which expansion is of the form $\sum_{n \in \mathbf{Z}_+} \mathrm{J}_n(f_n(\star, \cdot))$ and satisfies

$$\sum_{n \in \mathbf{Z}_+} (n+1)! \|\bar{f}_{n+1}\|_{\mathrm{L}^2(\mathbb{X})^{n+1}} < \infty,$$

and that is defined for any element $\mathrm{J}_n(f_{n+1}(\star, \cdot))$ of $\mathcal{U}$ by

$$\delta\big(\mathrm{J}_n(f_{n+1}(\star, \cdot))\big) := \mathrm{J}_{n+1}(\bar{f}_{n+1}), \tag{3.15}$$

where

$$\bar{f}_{n+1} = \frac{1}{n+1} \sum_{i=1}^{n+1} f_{n+1}\big((t_1, k_1), \cdots, (t_{i-1}, k_{i-1}), (t_{i+1}, k_{i+1}), \cdots, (t_{n+1}, k_{n+1}), (t_i, k_i)\big).$$

In the setting of classical Malliavin calculus, the divergence of adapted processes coincides with the Itô integral. We get the analogue in our context, where the role of the Itô integral is played by the $\mathcal{R}$-integral. Indeed, the equality $\delta(\mathbf{1}_\mathrm{A}) = \eta(\mathrm{A}) - \nu(\mathrm{A})$ holds for any $\mathrm{A} \in \mathcal{X}$ and leads for any $u \in \mathcal{U}$ to

$$\delta(u) = \mathrm{J}_1(u) = \sum_{(t,k) \in \mathbb{X}} \mathfrak{u}(\eta, (t,k)) \Delta \mathrm{R}_{(t,k)}. \tag{3.16}$$

This property holds for any $\mathbf{P} \otimes \nu$-square integrable process $u$. Let $u = \mathrm{J}_{n-1}(f_n(\star, \cdot))$ for some $f_n \in \mathrm{L}^2(\mathbb{X})^n$; the adaptedness of $u$ implies that $f_n(\star, (t,k)) = g_n(\star, (t,k)) \mathbf{1}_{\{1:t-1\}^n}$ for some $g_n \in \mathrm{L}^2(\mathbb{X})^n$. The result follows by writing

$$\delta(u) = \mathrm{J}_n(\bar{f}_{n+1}) = n \sum_{(t,k) \in \mathbb{X}} \mathrm{J}_{n+1}\big(\bar{g}_n(\star, (t,k)) \mathbf{1}_{\{1:t-1\}^n}\big) \Delta \mathrm{R}_{(t,k)} = \sum_{(t,k) \in \mathbb{X}} \mathfrak{u}(\eta, (t,k)) \Delta \mathrm{R}_{(t,k)}.$$

To state the property of closability of the gradient we need an integration by parts formula, appearing as a duality relation between D and $\delta$. Here is its version restricted to cylindrical functionals and simple processes.

**Proposition 3.8** (Integration by parts formula on $\mathcal{S} \times \mathcal{U}$). *For any* $(\mathrm{F}, u) \in \mathcal{S} \times \mathcal{U}$*,*

$$\mathbf{E}\left[\mathrm{F}\delta u\right] = \mathbf{E}\left[\langle \mathrm{DF}, u\rangle_{\mathrm{L}^2(\mathbb{X},\tilde{\nu})}\right]. \tag{3.17}$$

**Corollary 3.9** (Closability). *The operator* D *is closable from* $\mathrm{L}^2(\mathbf{P})$ *to* $\mathrm{L}^2(\mathbf{P} \otimes \nu)$.

By adjunction the operator $\delta$ is also closable from $\mathrm{L}^2(\mathbf{P} \otimes \nu)$ to $\mathrm{L}^2(\mathbf{P})$. Thus the domain $\mathbf{D}$ of D is the closure of $\mathcal{S}$ with respect to the norm

$$\|\mathrm{F}\|_{\mathbf{D}} := \left(\|\mathrm{F}\|^2_{\mathrm{L}^2(\mathbf{P})} + \|\mathrm{DF}\|^2_{\mathrm{L}^2(\mathbf{P}\otimes\tilde{\nu})}\right)^{1/2},$$

whereas the domain of $\delta$ is given by

$$\mathrm{dom}\ \delta = \left\{u \in \mathrm{L}^2(\mathbf{P} \otimes \nu) : \exists c > 0, \forall \mathrm{F} \in \mathbf{D},\ |\langle \mathrm{DF}, u\rangle|_{\mathrm{L}^2(\mathbf{P}\otimes\nu)} \leqslant c\|\mathrm{F}\|_{\mathrm{L}^2(\mathbf{P})}\right\}.$$

The integration by parts formula can be thus extended to the respective domains of D and $\delta$ to get what appears as a generalised commutation property: for any $\mathrm{F} \in \mathbf{D}$, $u \in \mathrm{dom}\ \delta$,

$$\mathbf{E}\left[\mathrm{F}\,\delta u\right] = \mathbf{E}\left[\langle \mathrm{DF}, u\rangle_{\mathrm{L}^2(\mathbb{X},\tilde{\nu})}\right]. \tag{3.18}$$

### 3.2.3 The Ornstein-Uhlenbeck structure

This section is devoted to the construction of an Ornstein-Uhlenbeck structure around the eponymous semi-group $(\mathrm{P}_\tau)_{\tau\in\mathbf{R}_+}$ and its generator L. Define the *Ornstein-Uhlenbeck semi-group* by its action on the chaotic decomposition: for any $\mathrm{F} \in \mathrm{L}^2(\mathbf{P})$ decomposed as (2.14),

$$\mathrm{P}_\tau \mathrm{F} = \sum_{n\in\mathbf{Z}_+} e^{-n\tau}\mathrm{J}_n(f_n).$$

**Proposition 3.10.** *The domain of the Ornstein-Uhlenbeck operator* L *(also called number operator) is the set of random variables* $\mathrm{F} \in \mathrm{L}^2(\mathbf{P})$ *which chaotic decomposition satisfies* (3.12) *(in particular* $\mathrm{dom\ L} \subset \mathbf{D}_0$*). For any* $\mathrm{F} \in \mathrm{dom\ L}$ *of expansion* (2.14)*,*

$$\mathrm{LF} = -\sum_{n\in\mathbf{N}} n\mathrm{J}_n(f_n).$$

*It satisfies the remarkable identity:* $\mathrm{F} \in \mathrm{dom\ L}$ *if and only if* $\mathrm{F} \in \mathbf{D}$ *and* $\mathrm{DF} \in \mathrm{dom}\ \delta$ *and, in this case,* $\mathrm{LF} = -\delta\mathrm{DF}$.

*Proof.* The identity $\mathrm{LF} = -\delta\mathrm{DF}$ can be stated first for $\mathrm{F} = \mathrm{J}_n(f_n)$ with $f_n \in \mathrm{L}^2(\mathbb{X})^n$, using (3.16) and then extended to $\mathbf{D}$ by closability of the operator D. □

The inverse of the number operator, denoted $\mathrm{L}^{-1}$ is defined on the subspace of $\mathrm{L}^2(\mathbf{P})$ made of random variables with null expectation, and that is given, for any F written as (2.11), by

$$\mathrm{L}^{-1}\mathrm{F} = -\sum_{n\in\mathbf{N}} \frac{1}{n}\mathrm{J}_n(f_n). \tag{3.19}$$

### 3.3 From combination of $\mathrm{L}^1$ and $\mathrm{L}^2$ theories to a unified Markov-Malliavin structure

As seen in the previous subsections, the operators $\mathrm{D}^+/\mathrm{D}$, $\widetilde{\delta}/\delta$ and $\widetilde{\mathrm{L}}/\mathrm{L}$, have a meaning either in $\mathrm{L}^1$ or $\mathrm{L}^2$ context. In this section we combine $\mathrm{L}^1$ and $\mathrm{L}^2$ theories to formalize a unified Markov-Malliavin structure.

#### 3.3.1 Operators $\mathrm{D}^+$ and $\mathrm{D}$: Stroock's formula

Within additional hypotheses, the operators $\mathrm{D}^+$ and $\mathrm{D}$ coincide. The very definition of the domain of the operator $\mathrm{D}$ and the chaotic decomposition ensure that if $\mathrm{F} \in \mathbf{D}$, then $\mathrm{DF} \in \mathrm{L}^2(\mathbf{P})$. The following lemma provides the reciprocal, as well as a more tractable expression of the Malliavin derivative, in terms of a difference operator acting on $\mathrm{L}^2(\mathbf{P})$.

**Proposition 3.11.** *Let* $\mathrm{F} \in \mathrm{L}^2(\mathbf{P})$. *If* $\mathrm{D}^+\mathrm{F} \in \mathrm{L}^2(\mathbf{P} \otimes \nu)$, *then* $\mathrm{F} \in \mathbf{D}$. *Moreover,*

$$\mathrm{DF} = \mathrm{D}^+\mathrm{F} \ ; \ \ \mathbf{P} \otimes \nu\text{−a.s.} \tag{3.20}$$

**Remark 3.12.** We can retrieve one of the specific identities existing in the Gaussian and Poisson spaces: for any process $u \in \mathcal{U}$, $\mathrm{D}_{(t,k)}\mathrm{J}_1(h) = h(t,k)$. Let $\mathfrak{r}_{(t,k)}$ be the representative of the process R. Applying (3.13) to $\mathrm{F} = \mathrm{J}_1(h)$ gives

$$\mathrm{D}_{(t,k)}\mathrm{J}_1(h) = \sum_{s \in \mathbf{N}} \sum_{\ell \in \mathrm{E}} h(s,\ell) \left[\mathfrak{r}(\pi_t(\eta) + \delta_{(t,k)}, (s,\ell)) - \mathfrak{r}(\eta, (s,\ell))\right] = h(t,k),$$

by noting that $\mathfrak{r}(\pi_t(\eta) + \delta_{(t,k)}, (s,\ell)) - \mathfrak{r}(\pi_t(\eta), (s,\ell)) = \mathbf{1}_{\{(t,k)\}}((s,\ell))$.

The integrands of multiple integrals appearing in the chaotic decomposition of F can be expressed in terms of iterative Malliavin derivatives of F. This entails the useful following lemmas. In fact, the operator $\mathrm{D}^+$ can be canonically iterated by letting $\mathrm{D}^{(1)} = \mathrm{D}^+$ and defining the $n$-th ($n \in \mathbf{N}$) *difference operator* by the recursion formula $\mathrm{D}^{(n)} = \mathrm{D}^+(\mathrm{D}^{(n-1)})$. We get explicitly for any $\mathrm{F} \in \mathcal{L}^0(\Omega)$,

$$\mathrm{D}^{(n)}_{(\mathbf{t}_n,\mathbf{k}_n)}\mathrm{F} = \mathrm{D}^+_{(t_1,k_1)}\big(\mathrm{D}^{(n-1)}_{(t_2,k_2),\dots,(t_n,k_n)}\mathrm{F}\big) = \sum_{\mathrm{J} \subset [n]} (-1)^{n-|\mathrm{J}|}\mathrm{F}\Big(\pi^{[n]}(\cdot) + \sum_{j \in \mathrm{J}} \delta_{(t_j,k_j)}\Big),$$

where $\pi^{[n]} := \bigcirc_{t=1}^{n} \pi_t$. This satisfies the remarkable identities that lead to the expression of the functions $f_n$ in (2.14) in terms of the $n$-th difference operator, called *Stroock's formula.*

**Lemma 3.13.** *For any* $\mathrm{F} \in \mathrm{L}^2(\mathbf{P})$,

$$\mathbf{E}\left[\mathrm{D}^{(n)}_{(\mathbf{t}_n,\mathbf{k}_n)}\mathrm{F}\right] = \mathbf{E}\left[\mathrm{F}\prod_{i=1}^{n} \frac{\Delta \mathrm{R}_{(t_i,k_i)}}{\kappa_i}\right].$$

**Lemma 3.14.** *For any* $\mathrm{F}, \mathrm{G} \in \mathrm{L}^2(\mathbf{P})$,

$$\mathbf{E}[\mathrm{FG}] = \mathbf{E}[\mathrm{F}]\,\mathbf{E}[\mathrm{G}] + \sum_{n \in \mathbf{N}} \frac{1}{n!} \langle \mathbf{E}[\mathrm{D}^{(n)}\mathrm{F}], \mathbf{E}[\mathrm{D}^{(n)}\mathrm{G}]\rangle_{\mathrm{L}^2(\mathbb{X},\nu)^{\otimes n}}.$$

**Proposition 3.15** (Stroock's formula)**.** *Let* $\mathrm{F} \in \mathrm{L}^2(\mathbf{P})$. *Then,* $\mathrm{D}^{(n)}\mathrm{F} \in \mathrm{L}^2(\mathbf{P} \otimes \nu^{\otimes n})$ *for any* $n \in \mathbf{N}$, *and* F *admits a chaotic decomposition of the form* (2.11) *with* $f_0 = \mathbf{E}[\mathrm{F}]$ *and*

$$f_n((\mathbf{t}_n,\mathbf{k}_n)) = \frac{1}{n!}\mathbf{E}\left[\mathrm{D}^{(n)}_{(\mathbf{t}_n,\mathbf{k}_n)}\mathrm{F}\right] \ ; \ \forall (\mathbf{t}_n,\mathbf{k}_n) \in \mathbb{X}^n. \tag{3.21}$$

### 3.3.2 $\widetilde{\mathrm{L}}$ and Ornstein-Uhlenbeck operators: Mehler's formula

In this part, we provide an integral representation of $(\mathrm{P}_\tau)_{\tau\in\mathbf{R}_+}$ in $\mathrm{L}^1(\mathbf{P})$, called *Mehler's formula*. We proceed in a similar fashion as done in the Poisson setting by Last, Peccati and Schulte [36]. Let $\eta \in \mathfrak{N}_{\mathbb{X}}$. Consider the binomial process N associated to $\eta$ and split it into two processes $\mathrm{N}^{(\gamma)}$ according to independent random draws of a Bernoulli random variable of mean $\gamma$. This means that any point charged by $\eta$ belongs to $\eta^{(\gamma)}$ with probability $\gamma$ and to $\eta^{(1-\gamma)}$ with probability $1-\gamma$. Important point: since the measure $\nu$ is not diffuse, we need to ensure that a point $(t,k)\in\eta$ cannot be simultaneously charged by $\eta^{(\gamma)}$ and $\eta^{(1-\gamma)}$. In other words, $(t,k)$ is either in the support of $\eta^{(\gamma)}$ or in that of $\eta^{(1-\gamma)}$. Considering $\eta$ as a proper process via Definition (2.1), let $\eta_{\mathrm{K}^\tau}$ be the $\mathrm{K}^\tau$-*marking* (see Last and Penrose [38], definition 5.3) of $\eta$ defined by

$$\eta_{\mathrm{K}^\tau} = \sum_{t=1}^{\mathrm{N}} \delta_{((\mathrm{T}_t,\mathrm{V}_t),\varepsilon_t^\tau)},$$

where $(\varepsilon_t^\tau)_{t\in\mathbf{N}}$ is a sequence of variables which conditional distribution given $\{\mathrm{N}=n\}$ (for $n\in\mathbf{N}$) **and** $\{(\mathrm{T}_t,\mathrm{V}_t),\, t\in\{1:n\}\}$, is that of independent random variables defined by $\varepsilon_t^\tau = \mathbf{1}_{\{\theta_t\leqslant\tau\}}$, and $(\theta_t)_{t\in\mathbf{N}}$ is a sequence of independent exponential random variables of mean 1. We can prove that $\eta_{\mathrm{K}^\tau}$ is a binomial process on $\mathbb{X}\times\{0,1\}$ of intensity measure $\nu\otimes\mathrm{K}^\tau$. Denote also

$$\eta^{\tau,0} := \eta_{\mathrm{K}^\tau}(\cdot\times\{0\}) \quad\text{and}\quad \eta^{\tau,1} := \eta_{\mathrm{K}^\tau}(\cdot\times\{1\}), \tag{3.22}$$

that are (not independent) binomial processes with respective intensities $e^{-\tau}\nu$ and $(1-e^{-\tau})\nu$. To see it, one can use the Laplace characterisation of binomial processes, that can be found in Last and Penrose (see [38], exercise 3.5): the Laplace transform of a mixed binomial process (which definition is given by (6.2)) with mixing measure $\mathbf{K}$ and sampling distribution $\mathbf{Q}$ is the function defined on $\mathbf{R}_+(\mathbb{X})$, the set of measurable functions from $\mathbb{X}$ to $\mathbf{R}_+$, by

$$\mathcal{L}_\eta(f) = \mathcal{G}_{\mathbf{K}}\Big(\int e^{-f}\,\mathrm{d}\mathbf{Q}\Big)\ ;\ f\in\mathbf{R}_+(\mathbb{X}),$$

where $\mathcal{G}_{\mathbf{K}}(x) := \sum_{n\in\mathbf{Z}_+}\mathbf{K}(\{n\})x^n$, for $x\in[0,1]$. Then, for any $f\in\mathbf{R}_+(\mathbb{X})$,

$$\mathcal{L}_{\eta^{\tau,0}}(f) = \sum_{n\in\mathbf{Z}_+}\mathbf{K}(\{n\})\Big(\sum_{k\in\mathrm{E}} e^{-f(k,0)}\,e^{-\tau}\mathbf{Q}(\{k\})\Big)^n = \mathcal{G}_{\mathbf{K}}\Big(\int_{\mathrm{E}\times\{0,1\}} e^{-f}\,(e^{-\tau}\mathrm{d}\delta_0\otimes\mathrm{d}\mathbf{Q})\Big),$$

so that $\eta^{\tau,0}$ is a binomial process with intensity $e^{-\tau}\nu$. The computation of the Laplace transform of $\eta^{\tau,0}+\eta^{\tau,1}$ suffices to see that the two processes are not independent. Nevertheless, we have $\eta^{\tau,0}+\eta^{\tau,1}=\eta$. The formula below is very similar to the one existing in the Poisson case that can be found in the work of Last, Peccati and Schulte [36] or in its original formulation in Privault (see [56], Lemma 6.8.1). The main difference lies in the presence here of the random variable $\varepsilon$. Implicitly defined in the thinning appearing in Mehler's formula for Poisson processes, it is explicitly required here to guarantee that a same point can not be weighted simultaneously by $\eta^{\tau,0}$ and $\tilde{\eta}$.

**Proposition 3.16.** *Let $\eta\in\widehat{\mathfrak{N}}_{\mathbb{X}}$ and $\mathrm{F}\in\mathrm{L}^1(\mathbf{P})$ of representative $\mathfrak{f}$. For any $\tau\in\mathbf{R}_+$,*

$$\mathrm{P}_\tau\mathrm{F} = \mathrm{P}_\tau\mathfrak{f}(\eta^{\tau,0}+\eta^{\tau,1}) = \int \mathbf{E}\left[\mathfrak{f}(\eta^{\tau,0}+\varepsilon^\tau\tilde{\eta})\middle|\eta\right]\Pi_\nu(\mathrm{d}\tilde{\eta})\ ;\ \mathbf{P}-\text{a.s.}, \tag{3.23}$$

*where $\Pi_\nu$ denotes the distribution of a marked binomial process of intensity measure $\nu$ and $\tilde{\eta}$ is a point process which distribution given $\eta$ follows the rule:*

$$\mathbf{P}((t,k) \in \tilde{\eta} \,|\, (t,k) \notin \eta) = \lambda \mathbf{Q}(\{k\}) \quad \textit{and} \quad \mathbf{P}((t,k) \notin \tilde{\eta} \,|\, (t,k) \in \eta) = 1 - \lambda \mathbf{Q}(\{k\}). \tag{3.24}$$

The first equality in (3.23) ensures that for any $\mathrm{F} \in \mathrm{L}^1(\mathbf{P})$, $\tau \in \mathbf{R}_+$,

$$\mathbf{E}\left[\mathrm{P}_\tau \mathrm{F}\right] = \mathbf{E}\left[\mathrm{F}\right],$$

while Jensen's inequality together with (3.23) imply the contractivity property of the semi-group: for any $p \in \mathbf{N}$,

$$\mathbf{E}\left[|\mathrm{P}_\tau \mathrm{F}|^p\right] \leqslant \mathbf{E}\left[|\mathrm{F}|^p\right]. \tag{3.25}$$

The semi-group $(\mathrm{P}_\tau)_{\tau \in \mathbf{R}_+}$ satisfies the usual commutation property:

**Proposition 3.17.** *For any $\mathrm{F} \in \mathrm{L}^2(\mathbf{P})$, and $\tau \in \mathbf{R}_+$,*

$$\mathrm{D}\mathrm{P}_\tau \mathrm{F} = e^{-\tau} \mathrm{P}_\tau \mathrm{D}\mathrm{F}. \tag{3.26}$$

The commutation property induces several and useful corollary results gathered in the following statement. Unfortunately, even if $\mathrm{F} \in \operatorname{dom} \mathrm{L}$ such that $\mathrm{DF} \in \mathrm{L}^1(\mathbf{P})$, we can not state $\mathrm{LF} = \widetilde{\mathrm{L}}\mathrm{F}$ $\mathbf{P}$-almost surely. Indeed, on the one hand, $\delta(\mathrm{DF}) = \sum_{(t,k)\in\mathbb{X}} (\mathrm{D}_{(t,k)}\mathrm{F}) \Delta \mathrm{R}_{(t,k)}$ whereas if $\mathrm{DF} \in \mathrm{L}^1(\mathbf{P})$, $\tilde{\delta}(\mathrm{DF}) = \sum_{(t,k)\in\mathbb{X}} (\mathrm{D}_{(t,k)}\mathrm{F}) \Delta \mathrm{Z}_{(t,k)}$. Nevertheless, follows from (2.18) that

$$\mathrm{LF} = -\sum_{(t,k)\in\mathbb{X}} (\mathrm{D}_{(t,k)}\mathrm{F}) \sum_{\ell \in \mathrm{E}} \mathfrak{m}_{k\ell}^{-1} \Delta \mathrm{Z}_{(t,\ell)} = \sum_{(t,\ell)\in\mathbb{X}} \sum_{k\in\mathrm{E}} \mathfrak{m}_{k\ell}^{-1} (\mathrm{D}_{(t,k)}\mathrm{F}) \Delta \mathrm{Z}_{(t,\ell)} =: \widetilde{\mathrm{L}}\widetilde{\mathrm{F}}, \tag{3.27}$$

where $\widetilde{\mathrm{F}}$ is a square-integrable random variable such that $\mathrm{D}_{(t,\ell)}\widetilde{\mathrm{F}} = \sum_{k\in\mathrm{E}} \mathfrak{m}_{k\ell}^{-1} (\mathrm{D}^+_{(t,k)}\mathrm{F})$ for any $(t,\ell) \in \mathbb{X}$. This is well and uniquely defined provided $\mathbf{E}[\widetilde{\mathrm{F}}]$ is given; indeed as a consequence of Clark formula (see forthcoming section 4), the knowledge of $(\mathrm{D}_{(t,k)}\mathrm{F},\, (t,k) \in \mathbb{X})$ and $\mathbf{E}\left[\mathrm{F}\right]$ provides the expression of F $\mathbf{P}$-almost surely.

**Corollary 3.18.** *For any $\mathrm{F} \in \mathrm{L}^2(\mathbf{P})$ such that $\mathbf{E}\left[\mathrm{F}\right] = 0$,*

$$\mathrm{L}^{-1}\mathrm{F} = -\int_0^\infty \mathrm{P}_\tau \mathrm{F} \,\mathrm{d}\tau, \qquad \mathbf{P} \otimes \nu - \text{a.e.} \tag{3.28}$$

*Moreover,*

$$-\mathrm{D}\mathrm{L}^{-1}\mathrm{F} = \int_0^\infty e^{-\tau} \mathrm{P}_\tau \mathrm{D}\mathrm{F} \,\mathrm{d}\tau, \qquad \mathbf{P} \otimes \nu - \text{a.e.} \tag{3.29}$$

**Remark 3.19.** The combination of Corollary 3.18 with the contraction property of $(\mathrm{P}_\tau)_{t\in\mathbf{R}_+}$ enables to bound $\mathrm{D}\mathrm{L}^{-1}\mathrm{F}$ with respect to the norm of DF: $\|\mathrm{D}\mathrm{L}^{-1}\mathrm{F}\|_{\mathrm{L}^2(\mathbf{P}\otimes\nu)} \leqslant \|\mathrm{DF}\|_{\mathrm{L}^2(\mathbf{P}\otimes\nu)}$.

**Remark 3.20.** In that case where E is a singleton, i.e. $\eta$ is a simple binomial process, we have $\Delta \mathrm{R}_t = \Delta \mathrm{Z}_t$ for any $t \in \mathbf{N}$, so that $\widetilde{\mathrm{L}} = \mathrm{L}$ and as a result $\widetilde{\Gamma} = \Gamma$ $\mathbf{P}$-almost surely (by letting $\Gamma(\mathrm{F},\mathrm{G}) = 1/2[\mathrm{L}(\mathrm{FG}) - \mathrm{F}(\mathrm{LG}) - \mathrm{G}(\mathrm{LF})]$, for $\mathrm{F}, \mathrm{G} \in \operatorname{dom} \mathrm{L}$). We retrieve thus the coincidence of $\mathrm{L}^1$ operators and Malliavin's ones within the remarkable association of $\mathrm{L}^1$ and $\mathrm{L}^2$ theories, as well as the natural link between Malliavin and Gamma calculus on the paradigm of $\mathrm{L}^2$ framework; both are highlighted in the Poisson case by Döbler and Peccati in [23].

As a conclusion, the combination of $L^1$ and $L^2$ theories is embodied by the existence of a $L^1$-correspondence $(D^+, \tilde{\delta}, \widetilde{L}, (P_\tau)_\tau)$ to the tuple of random objects $(D, \delta, L, (P_\tau)_\tau)$ that equips the space $(\mathfrak{N}_{\mathbb{X}}, \mathcal{F}, \mathbf{P})$. Moreover, each element of $(D, \delta, L, (P_\tau)_\tau)$ can be linked to one another of the tuple through one or a combination of the following identities and properties: the generalised integration by parts formula (3.18), the identity $L = -\delta D$, the generation of the semi-group $(P_\tau)_\tau$ by L, making it a Markov-Malliavin unified structure.

## 3.4 Comparison with pre-existing theories in the Poisson and Rademacher settings

It seems reasonable to ask: if we let the sequence V is deterministic constant equal to 1 (respectively $\lambda = 1$ and $E = \{-1, 1\}$), can we retrieve some element of stochastic analysis for Poisson processes on the real line (respectively Rademacher processes)?
Considering first the case where the sequence V be deterministic constant equal to 1 leads to define the orthogonal family $\mathcal{Z}^P = \{\Delta Z_t^P\, ; \, t \in \mathbf{N}\}$ by $\Delta Z_t^P = \mathbf{1}_{\{\Delta N_t = 1\}} - \lambda$, and the stochastic integral defined as the application $J_1^P : f \in L^2(\mathbf{N}, \lambda\#) \mapsto J_1^P(f)$ ($\#$ is the counting measure). The process $\eta$ is then the discrete analogue of the standard Poisson process on the real line. The gradient reads for $F \in \operatorname{dom} D^P$ of representative $\mathfrak{f}$,

$$D_t^P F = \mathfrak{f}(\pi_t(\eta) + \delta_t) - \mathfrak{f}(\pi_t(\eta)), \tag{3.30}$$

which is - up to a constant - a reminiscent of the gradient used by Decreusefond and Flint [19] on the Poisson space that is written (with corresponding notations): $\mathfrak{f}(\eta + \delta_t) - \mathfrak{f}(\eta - \delta_t)$. Nevertheless the operator $D^P$ is different from the usual one for Poisson processes on the real line: $\nabla_t F = \mathfrak{f}(\eta + \delta_t) - \mathfrak{f}(\eta)$. This definition is not suitable in the present context. Indeed, as stated by N. Privault (see [56], proof of the proposition 6.4.7),

$$\nabla_t J_n(f_n) = \mathbf{1}_{\{t \notin \eta\}} J_{n-1}(f_{n-1}(\star, t)), \tag{3.31}$$

which is $\mathbf{P}$-almost surely equal to $J_{n-1}(f_{n-1}(\star, t))$ since the intensity measure is diffuse in the Poisson case. This does not hold in our framework; the definition of the gradient (3.30) is thus justified to guarantee (3.31), in order to make the difference $(D^+)$ and annihilation (D) operators coincide and thereby combine $L^1$ and $L^2$ theories.

Let now $\lambda = 1$ and $E = \{-1, 1\}$. Basically, that means that the underlying binomial process jumps every time step. A Rademacher process $(X_t)_{t \in \mathbf{N}}$ can be defined by letting $X_t = V_t \in \{-1, 1\}$ and $Y_t := (2pq)^{-1/2}(\Delta Z_{(t,1)} + \Delta Z_{(t,-1)}) = (2pq)^{-1/2}(X_t - p + q)$ where the $\Delta Z_{t,\cdot}$ are defined as usual by (2.15) and $p := \mathbf{P}(X_t = 1) = 1 - q$. Thus $(Y_t)_{t \in \mathbf{N}}$ is a $\mathcal{F}$-(normal) martingale. By properly defining the function $g$ on $L^2(\mathbf{P})$ such that $F := F(X_1, \ldots, X_T) = g(\sqrt{2pq} Y_T)$ and $\overline{D}_t F := D_{(t,1)} g(\sqrt{2pq} Y_t) - D_{(t,-1)} g(\sqrt{2pq} Y_t)$, we get

$$\overline{D}_t Y_s = \frac{2}{\sqrt{2pq}} \mathbf{1}_{\{t\}}(s) = \sqrt{\frac{2}{pq}} \widehat{D} Y_s,$$

that is - up to a constant - the expression of the gradient $\widehat{D}$ defined on the Rademacher space (see for instance Privault [56], Proposition 1.6.2). All identities and formulas, such as the Clark formula and the predictable representation (see Privault [56], chapter 1), are inherited by construction.

# 4 Functional identities

In this section we derive some functional identities from our formalism. In fact, we can get the analogues of almost all identities existing in the Wiener or Poisson spaces that use the similar "Markov-Malliavin" structure. In a perspective of the forthcoming applications in the trinomial model we have chosen to focus on and present only two of them: Girsanov theorem and the Clark formula (and some corollaries).

## 4.1 Girsanov theorem

We provide our construction with the analogue of Girsanov theorem, which is reminiscent of that stated for compound Poisson processes (see Privault [58], Theorem 15.11).

**Theorem 4.1** (Girsanov theorem)**.** *Let $T \in \mathbf{N}$ and $\widetilde{\mathbf{P}}$ be a probability measure equivalent to $\mathbf{P}$ on $\mathcal{F}_T$. Then, there exist $\tilde{\lambda} \in (0,1)$ and a measure $\widetilde{\mathbf{Q}}$ on $\mathrm{E}$ such that $\widetilde{\mathbf{P}}$ is of compensator $\nu_{\widetilde{\mathbf{P}}} := \tilde{\lambda}\# \otimes \widetilde{\mathbf{Q}}$. Moreover, for any $t \in \{1:T\}$,*

$$\frac{\mathrm{d}\widetilde{\mathbf{P}}}{\mathrm{d}\mathbf{P}}\Big|_{\mathcal{F}_t} = \xi_t(h),$$

*where, if $\mathrm{E} = \{k^i,\, i \in \mathbf{Z}\}$, $h$ is the element of $\mathrm{L}^2(\mathbb{X})$ such that $\mathrm{J}_1(h) = \mathrm{J}_1(g\,;\mathcal{Z})$ with*

$$g(t,k^i) = \Big(\frac{\tilde{\lambda}\widetilde{\mathbf{Q}}(\{k^i\})}{\lambda\mathbf{Q}(\{k^i\})} - \frac{1-\tilde{\lambda}}{1-\lambda}\Big), \tag{4.1}$$

*for all $(t,i) \in \{1:T\} \times \mathbf{Z}$.*

**Corollary 4.2.** *Let $\tilde{\lambda} \in (0,1)$ and $\widetilde{\mathbf{Q}}$ be a measure on $\mathrm{E}$. Let $\varphi$ be the function defined on $\mathrm{E}$*

$$\varphi = \frac{\tilde{\lambda}(1-\lambda)}{\lambda(1-\tilde{\lambda})} \frac{\sum_{k\in\mathrm{E}} \widetilde{\mathbf{Q}}(\{k\})\, \mathbf{1}_{\{k\}}}{\sum_{k\in\mathrm{E}} \mathbf{Q}(\{k\})\, \mathbf{1}_{\{k\}}} - 1, \tag{4.2}$$

*Then, under the probability measure $\widetilde{\mathbf{P}}$ such that*

$$\mathrm{d}\widetilde{\mathbf{P}} = \left(\frac{1-\tilde{\lambda}}{1-\lambda}\right)^T \prod_{s=1}^{\mathrm{N}_T} (1+\varphi(\mathrm{V}_s))\, \mathrm{d}\mathbf{P},$$

*the process $\mathrm{Y}$ defined by (2.3) is a compound binomial process on $\mathbb{X}_T$ of intensity measure $\nu_{\widetilde{\mathbf{P}}} := \tilde{\lambda}\# \otimes \widetilde{\mathbf{Q}}$.*

**Remark 4.3.** The perturbations described by the *shift space* in Gaussian analysis (Cameron-Martin space for the Brownian motion in particular) act here on what characterizes the jumps: their occurrence and their height, respectively parametrised by $\lambda$ and $\mathbf{Q}$. The similar phenomenon is observed in the Poisson space.

## 4.2 Clark formula and corollaries

The Brownian martingale representation theorem says that a martingale adapted to the filtration of a Brownian motion is in fact a stochastic integral. The Clark formula gives the expression of the integrand of this stochastic integral in terms of the Malliavin gradient of the terminal value of the martingale. We here have the analogue formula.

**Proposition 4.4** (Clark formula)**.** *For any* $\mathrm{F} \in \mathrm{L}^2(\mathbf{P})$,

$$\mathrm{F} = \mathbf{E}\,[\mathrm{F}] + \sum_{(t,k)\in\mathbb{X}} \mathbf{E}\big[\mathrm{D}_{(t,k)}\mathrm{F} \,|\, \mathcal{F}_{t-1}\big]\, \Delta\mathrm{R}_{(t,k)}. \tag{4.3}$$

**Remark 4.5.** The operator $\mathrm{F} \in \mathrm{L}^2(\mathbf{P}) \mapsto \big(\mathbf{E}[\mathrm{D}_{(t,k)}\mathrm{F} \,|\, \mathcal{F}_{t-1}],\, (t,k) \in \mathbb{X}\big)$ is bounded with norm equal to 1. Indeed, from (4.3) together with the isometry property (2.9),

$$\big\|\mathbf{E}[\mathrm{D}_.\mathrm{F} \,|\, \mathcal{F}_{.-1}]\big\|_{\mathrm{L}^2(\mathbf{P}\otimes\tilde{\nu})} = \big\|\mathrm{F} - \mathbf{E}[\mathrm{F}]\big\|_{\mathrm{L}^2(\mathbf{P})} \leqslant \big\|\mathrm{F} - \mathbf{E}[\mathrm{F}]\big\|^2_{\mathrm{L}^2(\mathbf{P})} + \big(\mathbf{E}[\mathrm{F}]\big)^2 = \|\mathrm{F}\|^2_{\mathrm{L}^2(\mathbf{P})},$$

with equality in case $\mathrm{F} = \mathrm{J}_1(f_1)$ for some $f_1 \in \mathrm{L}^2(\mathbb{X})$.

As a direct consequence of Lemma 2.8 and Clark formula (4.3) we get the following corollary.

**Corollary 4.6.** *For any* $t \in \mathbf{N}$ *and* $\mathrm{F} \in \mathrm{L}^2(\mathbf{P})$,

$$\mathrm{F} = \mathbf{E}\big[\mathrm{F}|\mathcal{F}_t\big] + \sum_{s\geqslant t+1}\sum_{k\in\mathrm{E}} \mathbf{E}\big[\mathrm{D}_{(s,k)}\mathrm{F} \,|\, \mathcal{F}_{s-1}\big]\Delta\mathrm{R}_{(s,k)} \tag{4.4}$$

We can state the analogue of the so-called Chernoff-Nash-Poincaré inequality of Gaussian analysis (see Chernoff [14], Nash [42]). Our result is clearly a reminiscence of its counterpart in the Poisson space (see Last and Penrose [37], Wu [74]) or for independent random variables (see Decreusefond and Halconruy [20]).

**Corollary 4.7** (Poincaré inequality)**.** *For any* $\mathrm{F} \in \mathrm{L}^2(\mathbf{P})$,

$$\mathrm{var}(\mathrm{F}) \leqslant \mathbf{E}\left[\int_{\mathbb{X}} |\mathrm{D}_{(t,k)}\mathrm{F}|^2 \,\mathrm{d}\tilde{\nu}(t,k)\right].$$

**Remark 4.8.** Assume $\mathrm{E} = \{k^1, \cdots, k^n\}$. The transposition of the Clark formula in terms of $\mathcal{Z}$-integrals can be easily deduced from (4.4) together with (2.17); for any $\mathrm{F} \in \mathrm{L}^2(\mathbf{P})$,

$$\begin{aligned}\mathrm{F} = \mathbf{E}[\mathrm{F}] + \sum_{s\geqslant t+1}\sum_{j=1}^{n}\sum_{i=1}^{j} \mathfrak{m}^{-1}_{k^j k^i}\mathbf{E}\big[\mathrm{D}_{(s,k^j)}\mathrm{F} \,|\, \mathcal{F}_{s-1}\big]\Delta\mathrm{Z}_{(s,k^i)} \\ = \mathbf{E}[\mathrm{F}] + \sum_{s\geqslant t+1}\sum_{i=1}^{n} \mathbf{E}\big[\mathrm{D}^{\mathcal{Z}}_{(s,k^i)}\mathrm{F} \,|\, \mathcal{F}_{s-1}\big]\Delta\mathrm{Z}_{(s,k^i)},\end{aligned}$$

where $\mathrm{D}^{\mathcal{Z}}_{(s,k^i)}\mathrm{F} = \sum_{j=i}^{n} \mathfrak{m}^{-1}_{k^j k^i}\mathrm{D}_{(s,k^j)}\mathrm{F}$.

# 5 Applications

## 5.1 Malliavin-Stein method for compound Poisson approximation

The Stein method, initially developed to quantify the rate of convergence in the Central Limit Theorem (see Stein [71]) and then for Poisson convergence (see for instance Barbour, Lars and Janson [8]), has become a very popular not to say the most famous procedure to assess distances between two probability measures of the form

$$\mathrm{dist}_{\mathcal{T}}(\mathbf{P},\mathbf{Q}) = \sup_{h\in\mathcal{T}}\Big|\int_{\mathfrak{F}} h\,\mathrm{d}\mathbf{P} - \int_{\mathfrak{F}} h\,\mathrm{d}\mathbf{Q}\Big|, \tag{5.1}$$

where $\mathcal{T}$ is a class of real-valued test functions. The class $\mathcal{T}$ is furthermore separating, in the sense that if $\int_{\mathfrak{F}} h \, \mathrm{d}\mathbf{P} = \int_{\mathfrak{F}} h \, \mathrm{d}\mathbf{Q}$ for all $h \in \mathcal{T}$ if and only if $\mathbf{Q} = \mathbf{P}$. In particular, if $\mathcal{T} = \{\mathbf{1}_{\mathrm{A}}, \mathrm{A} \in \mathcal{X}\}$ coincides with the total *total-variation distance* and will be denoted $\mathrm{dist}_{\mathrm{TV}}$. The first one consists in converting the difficult initial problem (5.1) by the more tractable expression

$$\sup_{\varphi \in \mathcal{K}} \Big| \mathbf{E}\left[\mathrm{L}\varphi(\mathrm{Y})\right] \Big| = \sup_{\varphi \in \mathcal{K}} \Big| \mathbf{E}\left[\mathrm{L}_1\varphi(\mathrm{Y}) + \mathrm{L}_2\varphi(\mathrm{Y})\right] \Big|,$$

where Y is a random variable of law $\mathbf{Q}$, L and $\mathcal{K}$ are respectively the Stein operator and the Stein class associated to the target measure $\mathbf{P}$. The aim of the second step is to developp tools in order to transform $\mathrm{L}_1\varphi(\mathrm{X})$ into $-\mathrm{L}_2\varphi(\mathrm{X}) + \text{remainder}$. This remainder is what gives the bound of the distance and, in a problem of convergence, provides its rate. Besides, Nourdin and Peccati showed in [43] that this transformation step can be advantageously performed using integration by parts in the sense of Malliavin calculus. In this section, we make use of our formalism to provide an analogue of the Stein-Malliavin criterion for the Poisson (respectively compound Poisson) approximation by binomial (respectively marked binomial) functionals with respect to the total variation distance. This is defined for two $\mathbf{Z}_+$-random variables X and Y (the case of interest here), by

$$\mathrm{dist}_{\mathrm{TV}}(\mathbf{P}_{\mathrm{X}}, \mathbf{P}_{\mathrm{Y}}) = \sup_{\mathrm{A} \subset \mathbf{Z}_+} \big| \mathbf{P}(\mathrm{X} \in \mathrm{A}) - \mathbf{P}(\mathrm{Y} \in \mathrm{A}) \big|.$$

First, we can state a result for the Poisson approximation, in the same spirit as Peccati [53]. Within the same framework, let $\mathcal{P}(\lambda_0)$ the Poisson law with parameter $\lambda_0$. Consider for a given function $\varphi : \mathbf{Z}_+ \to \mathbf{R}$, $\nabla\varphi$ the *forward difference* $\nabla\varphi := \varphi(\cdot + 1) - \varphi$, and $\nabla^2\varphi$ its second iteration $\nabla^2 := \nabla(\nabla\varphi)$, that satisfies the useful (as proved in particular in Peccati [53], proof of Theorem 3.3) inequality: for all $a, k \in \mathbf{Z}_+$,

$$\big|\varphi(k) - \varphi(a) - \nabla\varphi(a)(k-a)\big| \leqslant \frac{\|\nabla^2\varphi\|}{2} |(k-a)(k-a-1)|. \tag{5.2}$$

For any $\mathrm{A} \subset \mathbf{Z}_+$, we denote by $\varphi_{\mathrm{A}} : \mathbf{Z}_+ \to \mathbf{R}$ the unique solution to the Chen-Stein equation

$$\mathbf{P}(\mathcal{P}(\lambda_0)) - \mathbf{1}_{\mathrm{A}}(k) = k\varphi_{\mathrm{A}}(k) - \lambda_0\varphi_{\mathrm{A}}(k+1) \; ; \; k \in \mathbf{Z}_+, \tag{5.3}$$

satisfying the boundary condition $\nabla^2\varphi_{\mathrm{A}}(0) = 0$. The function class $\mathcal{K} = \{\varphi_{\mathrm{A}}, \ \mathrm{A} \subset \mathbf{Z}_+\}$ fulfils the estimates (be also found for instance in Peccati [53]),

$$\|\varphi\|_\infty \leqslant \min\Big(1, \sqrt{\frac{2}{e\lambda_0}}\Big), \ \|\nabla\varphi\|_\infty \leqslant \frac{1 - e^{-\lambda_0}}{\lambda_0}, \text{ and } \|\nabla^2\varphi\|_\infty \leqslant \frac{2 - 2e^{-\lambda_0}}{\lambda_0^2},$$

where we have denoted $\|\varphi\|_\infty = \max_{\mathrm{A} \subset \mathbf{Z}_+} \|\nabla\varphi_{\mathrm{A}}\|_\infty$, $\|\nabla\varphi\|_\infty = \max_{\mathrm{A} \subset \mathbf{Z}_+} \|\nabla\varphi_{\mathrm{A}}\|_\infty$ and $\|\nabla^2\varphi\|_\infty = \max_{\mathrm{A} \subset \mathbf{Z}_+} \|\nabla^2\varphi_{\mathrm{A}}\|_\infty$.

**Theorem 5.1.** *Consider* $\lambda_0 \in \mathbf{R}_+^*$ *and let* F *be a square-integrable* $\mathbf{Z}_+$*-valued random variable such that* $\mathbf{E}[\mathrm{F}] = \lambda_0$. *Then,*

$$\begin{aligned}\mathrm{dist}_{\mathrm{TV}}\big(\mathbf{P}_{\mathrm{F}}, \mathcal{P}(\lambda_0)\big) \leqslant \frac{1 - e^{-\lambda_0}}{\lambda_0} \mathbf{E}\Big[\big|\lambda_0 - \langle \widetilde{\mathrm{D}}\mathrm{F}, -\mathrm{D}\mathrm{L}^{-1}(\mathrm{F} - \mathbf{E}[\mathrm{F}])\rangle_{\mathrm{L}^2(\mathbb{X},\nu)}\big)\big|\Big] \\ + \frac{1 - e^{-\lambda_0}}{\lambda_0^2} \mathbf{E}\Big[\int_{\mathbf{N}} \big|(\widetilde{\mathrm{D}}_t\mathrm{F})(\widetilde{\mathrm{D}}_t\mathrm{F} - 1)\big|\big|\mathrm{D}_t\mathrm{L}^{-1}(\mathrm{F} - \mathbf{E}[\mathrm{F}])\big| \, \nu(\mathrm{d}t)\Big].\end{aligned} \tag{5.4}$$

The aim is now to provide such a bound for the compound Poisson approximation. Let $\mathcal{PC}(\lambda_0, \mathbf{V})$ denote the law of a compound Poisson variable of parameters $(\lambda_0, \mathbf{V})$, that means it can be written as the distribution of the variable

$$\sum_{i=1}^{\mathrm{N}^{\mathrm{P}}} \mathrm{V}_i,$$

where $\mathrm{N}^{\mathrm{P}}$ is a Poisson random variable of mean $\lambda_0$ and $\{\mathrm{V}_i,\ i \in \mathbf{N}\}$ is a family of independent non-negative random variables of distribution $\mathbf{V}$. For any $\mathrm{A} \subset \mathbf{Z}_+$, denote $\psi_{\mathrm{A}}$ the unique solution of the Chen-Stein equation

$$\mathbf{1}_{\mathrm{A}}(\ell) - \mathbf{P}(\mathcal{PC}(\lambda_0, \mathbf{V})) = \ell \psi_{\mathrm{A}}(\ell) - \int_{\mathbb{X}} k \psi_{\mathrm{A}}(\ell + k) \, \mathrm{d}\nu(t, k) \ ; \ \ell \in \mathbf{Z}_+. \tag{5.5}$$

The function class $\mathcal{K}' = \{\varphi_{\mathrm{A}},\ \mathrm{A} \subset \mathbf{Z}_+\}$ satisfies the following estimate (see Erhardsson [25], Theorem 3.5)

$$\mathfrak{d}_{\mathcal{PC}} := \max_{\mathrm{A} \subset \mathbf{Z}_+} \|\psi_{\mathrm{A}}\|_\infty \vee \max_{\mathrm{A} \subset \mathbf{Z}_+} \|\nabla \psi_{\mathrm{A}}\|_\infty \leqslant \min\Big(1, \frac{1}{\lambda_0 \mathbf{V}(\{1\})}\Big) e^{\lambda_0}. \tag{5.6}$$

**Proposition 5.2.** *Consider* $\lambda_0 \in \mathbf{R}_+^*$ *and* $\mathbf{V}$ *a probability distribution on* $\mathbf{N}$*. Let* $\mathrm{V}_1$ *be a random variable of law* $\mathbf{V}$ *and* $\mathrm{F}$ *a square-integrable* $\mathbf{Z}_+$*-valued random variable such that* $\mathbf{E}[\mathrm{F}] = \lambda_0 \mathbf{E}[\mathrm{V}_1]$.

$$\begin{aligned} &\mathrm{dist}_{\mathrm{TV}}\big(\mathbf{P}_{\mathrm{F}}, \mathcal{PC}(\lambda_0, \mathbf{V})\big) \\ &\qquad \leqslant \Big| \int_{\mathbb{X}} \big[\mathrm{D}^+ \widetilde{\mathrm{L}}^{-1}(\mathfrak{f}(\eta) - \mathbf{E}[\mathfrak{f}(\eta)]) \psi_{\mathrm{A}}(\mathfrak{f}(\pi_t(\eta) + \delta_{(t,k)})) - k \psi_{\mathrm{A}}(\mathfrak{f}(\eta) + k)\big] \mathrm{d}\nu(t, k) \Big| \\ &\qquad\quad + \mathfrak{d}_{\mathcal{PC}} \Big| \int_{\mathbb{X}} \big[\mathrm{D}^+ \widetilde{\mathrm{L}}^{-1}(\mathfrak{f}(\eta) - \mathbf{E}[\mathfrak{f}(\eta)]) - k\big] \, \mathrm{d}\nu(t, k) \Big|. \end{aligned} \tag{5.7}$$

**Remark 5.3.** This result is only of interest in the case of the variable F is a marked binomial functional in the first chaos i.e. $\mathrm{F} = \mathrm{J}_1(f)$ for some $f \in \mathrm{L}^2(\mathbb{X})$ (which corresponds to the framework of the application of subsection 5.3) and is no relevant for more complicated functionals. In this latter, we can provide a bound by means of a Taylor expansion and in terms of the iterated operator $\nabla^2$. That turns out to be sub-optimal in the first chaos case we will be interested in later, which justifies our choice not to present it.

## 5.2 Head run problems

Consider a large number of independent throws of a coin of success (falling on face) probability $p \in (0, 1)$. Whatever the value of $p$, there will be sequences where the coin will fall on face each time; this sequence is called a *head run* and we aim at computing the probability that U, the lenght of the longest run of heads beginning in the first $n$ tosses, will be less to a test length $m \in \mathbf{N}$. The critical fact is that head runs occur in *clumps*; indeed, if there is the head of a run of length $m$ at position $i$, then with probability $p$, there will also be a run of length $m$ at position $i + 1$. We need then to "declump" the sequences in order to count only the first count. To do that, let $(\mathrm{C}_i)_{i \in \mathbf{N}}$ be a sequence of independent and identically distributed

Bernoulli variables of parameter $p$. Let $m$ be a fixed positive integer "test" value and consider the random variable

$$\mathrm{U} = \prod_{i=1}^{m} \mathrm{C}_i + \sum_{i=2}^{n} (1 - \mathrm{C}_{i-1}) \mathrm{C}_i \mathrm{C}_{i+1} \cdots \mathrm{C}_{i+m-1}$$

that gives the total number of clumps of runs of length $m$ or more. Note that $\mathbf{E}[\mathrm{U}] = p^m((n-1)(1-p)+1) =: \lambda_0$. Let N be a binomial process of intensity $p$. The random variable U can be rewritten as

$$\mathrm{U} = \prod_{i=1}^{m} \Delta \mathrm{N}_i + \sum_{i=1}^{n-1} (1 - \Delta \mathrm{N}_i) \Delta \mathrm{N}_{i+1} \Delta \mathrm{N}_{i+2} \cdots \Delta \mathrm{N}_{i+m} =: \sum_{i=0}^{n-1} \mathrm{U}_i, \tag{5.8}$$

so that it appears as a binomial functional with mean $p$ since the sequence V is here deterministic and constant equal to 1. Its chaotic decomposition reads

$$\mathrm{U} = \mathbf{E}[\mathrm{U}] + \sum_{j=1}^{m+1} \mathrm{J}_j\big(f_j \mathbf{1}_{[n+m-1]}\big),$$

where, in particular, for $j \in \{1 : m+1\}$, $f_j(t_1, t_2, \cdots, t_j) = 0$, as soon as $\prod_{i=1}^{j-1} \mathbf{1}_{\{1\}}(t_{i+1} - t_i) = 0$. Since $\mathrm{U} \in \mathbf{D}$, $\mathrm{DU} = \mathrm{D}^+\mathrm{U}$ $\mathbf{P} \otimes \nu$-almost surely, and

$$\mathrm{D}_t\mathrm{U} = \prod_{i=1, i\neq t}^{m} \Delta \mathrm{N}_i + \sum_{i=1}^{n-1} \Big( \mathbf{1}_{[i+1, i+m]}(t)(1 - \Delta \mathrm{N}_i) \prod_{\ell=1, i+\ell\neq t}^{m} \Delta \mathrm{N}_{i+\ell} + \mathbf{1}_{\{i\}}(t) \prod_{\ell=1}^{m} \Delta \mathrm{N}_{i+\ell} \Big).$$

**Theorem 5.4.** *Let $\lambda_0 = p^m((n-1)(1-p)+1)$. Then,*

$$\mathrm{dist}_{\mathrm{TV}}(\mathbf{P}_{\mathrm{U}}, \mathcal{P}(\lambda_0)) \leqslant p^{2m}[2(m-1)q^2 + 2mq + 1] + (n-m+1)(1-p)p^{2m+1} \|\nabla \varphi\|_\infty.$$

**Remark 5.5.** The previous result gives an insight into the distribution of $\mathrm{T}_n$, the length of the longest head run. As explained in Arratia, Goldstein and Gordon [2], as a consequence of the previous theorem, the distribution of $\mathrm{T}_n$ may be approximated as

$$\mathbf{P}(\mathrm{T}_n < t) = \mathbf{P}(\mathrm{U} = 0) = e^{-\lambda_0}.$$

The definition of a test lenght requires that $\lambda_0$ is bounded away from 0 and $\infty$. In other words, this means the existence of a deterministic constant $c$ such that

$$m = \log_{1/p}\big((n-1)(1-p)+1\big) + c.$$

This assumption entails that the Poisson approximation in Theorem 5.4 is of order $1/n$ where the constant can be found by considering the $1/n$-order terms such as $np^{2m}$. We thus find the result of Arratia, Goldstein and Gordon in [2] who used the Chen-Stein method to deal with the local dependence structure of U.

## 5.3 Number of occurrences of a word in a DNA sequence

A DNA sequence can be represented by a finite series $\mathrm{X}_1\mathrm{X}_2\ldots\mathrm{X}_n$ of characters taken from the alphabet $\mathcal{A} := \{\mathrm{A},\mathrm{C},\mathrm{G},\mathrm{T}\}$ where the four letters stand for the four bases *adenine, cytosine, guanine and thymine.* The question of the identification of words W with unexpected frequencies is crucial in DNA sequence analysis, and in diagnostic issues in particular. In this example, we model the sequence $\mathrm{X}_1\mathrm{X}_2\ldots\mathrm{X}_n$ with an homogeneous and stationary Markov chain of order $m$. The transition probability is given by the application $\theta$ defined on $\mathcal{A}\times\mathcal{A}$, whereas the invariant probability measure is denoted by $\mu$. The aim is to compute the number of occurrences in the sequence of a given word W of size $h$ (with $h > m$) $\mathrm{W} = w_1w_2\cdots w_h$. Let $(\mathrm{Z}_j)_{j\in\mathcal{J}}$ be the sequence defined by

$$\mathrm{Z}_j = \mathbf{1}_{\{\mathrm{X}_j=w_j,\ldots,\mathrm{X}_{j+h-1}=w_h\}},$$

where $\mathcal{J} = \{1,\ldots,n-h+1\}$. Since the underlying Markov chain is homogeneous and stationary of invariant measure $\mu$, $\mathbf{E}[\mathrm{Z}_j] = \mu(\mathrm{W})$ $(j\in\mathcal{J})$. The number of occurrences of the word W is then provided by the random variable

$$\mathfrak{T}(\mathrm{W}) = \sum_{j\in\mathcal{J}} \mathrm{Z}_j,$$

whose asymptotic behaviour we want to analyse when $n$ goes to infinity and $h$ grows as $\log(n)$. As explained in particular in Schbath [66], the word W may appear in clumps. Indeed, if W has a periodic decomposition, its occurrences in the sequence can overlap. A $k$-clump is thus the occurrence of a concatenated word C composed of exactly $k$ overlapping occurrences of W. For instance, if W = ACTAA, the sequence

G<u>ACTAACTAAACTAA</u>TGAA<u>ACTAA</u>CG

has a 3-clump at position $j = 2$ and a 1-clump at position $j = 20$. Especially when the word W can overlap, we must consider $(\widetilde{\mathrm{Z}}_j)_{j\in\mathcal{J}}$, the "declumped" sequence associated to $(\mathrm{Z}_j)_{j\in\mathcal{J}}$, such that $\widetilde{\mathrm{Z}}_j$ only counts occurrences that do not overlap the preceding one. Define for any $j\in\mathcal{J}$,

$$\widetilde{\mathrm{Z}}_j = \mathrm{Z}_j(1-\mathrm{Z}_{j-1})\cdots(1-\mathrm{Z}_{j-h+1}).$$

**Remark 5.6.** In fact, as highlighted by Schbath ([66], remark 2) it would be more rigorous from a practical point of view to consider the "observable" sequence $(\widehat{\mathrm{Z}}_j)_{j\in\mathrm{J}}$ defined by $\widehat{\mathrm{Z}}_1 = \mathrm{Z}_1$ and for any $i\in\{2,\ldots,j-1\}$, $\widehat{\mathrm{Z}}_j = \mathrm{Z}_j\prod_{i=1}^{j}(1-\mathrm{Z}_i)$ and $\widehat{\mathrm{Z}}_j = \widetilde{\mathrm{Z}}_j$ otherwise, since $\mathrm{X}_0,\mathrm{X}_{-1},\ldots,\mathrm{X}_{-h+2}$ may not be known. That being so, as the total variation distance between $\widehat{\mathfrak{T}}(\mathrm{W}) = \sum_{j\in\mathcal{J}}\widehat{\mathrm{Z}}_j$ and $\widetilde{\mathfrak{T}}(\mathrm{W}) = \sum_{j\in\mathcal{J}}\widetilde{\mathrm{Z}}_j$ is bounded by $2h\mu(\mathrm{W})$, both distributions have the same asymptotic behaviour, so that the sequence $(\widetilde{\mathrm{Z}}_j)_{j\in\mathcal{J}}$ can be used more conveniently.

Define for any $k\in\mathbf{N}$, the random variable $\overline{\mathfrak{T}}^{(k)}(\mathrm{W})$ that gives the number of $k$-clumps, as well as for $(j,k)\in\mathcal{J}\times\mathbf{N}$, the random variable $\overline{\mathrm{Z}}_j^{(k)}$ that indicates if there is a $k$-clump at position $j$. In order to approximate $\mathfrak{T}(\mathrm{W})$, write up to now $\mathfrak{T}(\mathrm{W}) = \sum_{k\in\mathbf{N}} k\overline{\mathfrak{T}}^{(k)}(\mathrm{W})$ that can be well approximated (see for instance Barbour and Chryssaphinou [7], or Reinert and Schbath [60]) by the random variable

$$\overline{\mathfrak{T}}(\mathrm{W}) := \sum_{j\in\mathcal{J}}\sum_{k\in\mathbf{N}} k\overline{\mathrm{Z}}_j^{(k)},$$

and $\mathrm{dist}_{\mathrm{TV}}(\mathbf{P}_{\mathfrak{T}(\mathrm{W})}, \mathbf{P}_{\overline{\mathfrak{T}}(\mathrm{W})}) \leqslant 2h\mu(\mathrm{W})$. Moreover, it appears (see Robin [62] or Reinert and Schbath [60]) that for any $j \in \mathcal{J}$, $\overline{\mathrm{Z}}_j^{(k)}$ is a Bernoulli-distributed random variable of mean

$$p_k = (1-\alpha)^2 \alpha^{k-1} \mu(\mathrm{W}) \tag{5.9}$$

where $\alpha$ can be written with respect to the principal periods of W. The reader can find a explicit expression of $\alpha$ in the case of a first-order Markov chain in Schbath [66], section 3. This last point suggests to approximate $\mathfrak{T}(\mathrm{W})$ by $\overline{\mathfrak{T}}(\mathrm{W})$ and, in order to get a marked binomial functional, by introducing the random variable

$$\mathrm{H} = \sum_{j \in \mathcal{J}} \mathrm{V}_j \Delta \mathrm{N}_j,$$

where $(\mathrm{V}_j)_{j \in \mathbf{N}}$ is a sequence of independent and identically distributed random variables which the common geometric distribution $\mathbf{V}$ of parameter $(1-\alpha)$, where $\alpha$ appears in (5.9). In fact, $(1-\alpha)\alpha^{k-1}$ is the probability that the word W overlaps exactly $k$ times after having occurred at position $j$. The sequence V is also supposed to be independent of a Bernoulli process $(\Delta \mathrm{N}_j)_{j \in \mathcal{J}}$ of intensity $(1-\alpha)\mu(\mathrm{W})$ so that $\mathcal{PC}(\lambda_0, \mathbf{V})$ is exactly the Pólya-Aeppli distribution of parameters $(\lambda_0, \alpha)$ where $\lambda_0 = (n-h+1)(1-\alpha)\mu(\mathrm{W})$. Some computations highlight that $\overline{\mathfrak{T}}(\mathrm{W})$ and $\mathbf{P}_{\mathrm{H}}$ are identically distributed, so that $\mathrm{dist}_{\mathrm{TV}}(\mathbf{P}_{\overline{\mathfrak{T}}(\mathrm{W})}, \mathbf{P}_{\mathrm{H}}) = 0$. It remains to control $\mathrm{dist}_{\mathrm{TV}}\big(\mathbf{P}_{\mathrm{H}}, \mathcal{PC}(\lambda_0, \mathbf{V})\big)$ by means of Theorem 5.2. The following bound results from it.

**Proposition 5.7.** *Let $\lambda_0 = (n-h+1)(1-\alpha)\mu(\mathrm{W})$ and $\mathcal{PC}(\lambda_0, \mathbf{V})$ the random variable written as a Poisson of mean $\lambda_0$ compounded by the distribution $\mathbf{V}$. Then*

$$\begin{aligned}\mathrm{dist}_{\mathrm{TV}}\big(\mathbf{P}_{\mathfrak{T}(\mathrm{W})}, \mathcal{PC}(\lambda_0, \mathbf{V})\big) &\leqslant \mathrm{dist}_{\mathrm{TV}}(\mathbf{P}_{\mathfrak{T}(\mathrm{W})}, \mathbf{P}_{\overline{\mathfrak{T}}(\mathrm{W})}) + \mathrm{dist}_{\mathrm{TV}}(\mathbf{P}_{\overline{\mathfrak{T}}(\mathrm{W})}, \mathbf{P}_{\mathrm{H}}) \\ &\quad + \mathrm{dist}_{\mathrm{TV}}(\mathbf{P}_{\mathrm{H}}, \mathcal{PC}(\lambda_0, \mathbf{V})) \\ &\leqslant 2h\mu(\mathrm{W}) + (n-h+1)\mathfrak{d}_{\mathcal{PC}}\mu(\mathrm{W})^2\end{aligned}$$

*where $\mathfrak{d}_{\mathcal{PC}}$ is defined by* (5.6).

**Remark 5.8.** The convergence occurs since the assumption on the order of the length $h$ (in $\log n$) entails that $n\mu(\mathrm{W}) = \mathrm{O}(1)$ (see Schbath [66]). We retrieve the rate of convergence of this approximation in $\log n/n$ , without the additional assumptions made on the size of the "neighbourhood of dependence" (see Schbath [66]) or on the order of the magnitude of the maximal overlap (see Geske *et al.* [29]). As noted in several works, in particular Robin and Schbath [63], the compound Poisson approximation is an excellent choice (especially with respect to the Gaussian approximation and to a lesser extent to the Poisson one) to describe the asymptotic behavior of a long and rare word in an "infinite" DNA sequence.

## 5.4 Portfolio optimisation in the trinomial model

We consider a simple financial market modelled by two assets i.e. a couple of $\mathbf{R}_+$-valued processes $(\mathrm{A}_t, \mathrm{S}_t)_{t \in \mathbf{T}}$, defined on the same probability space $(\Omega, \mathcal{F}, \mathbf{P})$ where $\mathcal{F} = (\mathcal{F}_t)_{t \in \mathbf{T}}$ is a filtration (generally that generated by the canonical process) and $\mathbf{T} = \mathbf{Z}_+ \cap [0, T]$ ($T \in \mathbf{N}$) is called the trading interval. Denote also $\mathbf{T}^* = \mathbf{T} \setminus \{0\}$. The riskless asset $(\mathrm{A}_t)_{t \in \mathbf{T}}$ is deterministic with initial value $A_0 = a_0$ and is defined for $r \in \mathbf{R}_+$ (generally smaller than 1) by

$$A_t = a_0(1+r)^t, \tag{5.10}$$

whereas the stock price which models the risky asset, is the $\mathcal{F}$-adapted process $(S_t)_{t\in\mathbf{T}}$ with (deterministic) initial value $S_0 = 1$ and such that for any $t \in \mathbf{T}^*$,

$$\Delta S_t = \eta_t \, S_{t-1} \, \Delta N_t, \tag{5.11}$$

where $\eta_t = b\mathbf{1}_{\{W_t=1\}} + a\mathbf{1}_{\{W_t=-1\}}$, $a$ and $b$ are real numbers such that $-1 < a < r < b$ and $\{W_t,\, t \in \mathbf{T}^*\}$ is a family of i.i.d. $\{-1,1\}$-valued random variables such that $\mathbf{P}(W_t = 1) = p$ ($p \in (0,1)$, $q = 1-p$). The sequence of discounted prices $(\overline{S}_t)_{t\in\mathbf{T}}$ is defined by $\overline{S}_t = A_t^{-1} S_t$ ($t \in \mathbf{T}$).

**Remark 5.9** (Trinomial and ternary models: differences and equivalence in law)**.** The price process defined in the ternary model has the same law as the one of a well-chosen trinomial model (for more details on classical trinomial model see for instance Delbaen [22] or Runggaldier [64]). As a reminder, the stock price $(T_t)_{t\in\mathbf{T}}$ is defined in this latter model by $T_0 = 1$ and verifies the recurrent relation

$$T_t = (1+b)T_{t-1}\mathbf{1}_{\{X_t=1\}} + (1+a)T_{t-1}\mathbf{1}_{\{X_t=-1\}} + T_{t-1}\mathbf{1}_{\{X_t=0\}},$$

where the process $(X_t)_{t\in\mathbf{T}^*}$ is distributed according to the measure $\mathbf{P}$ such that

$$\mathbf{P}(X_t = 1) = \bar{p},\ \mathbf{P}(X_t = -1) = \bar{q} \ \text{ and } \mathbf{P}(X_t = 0) = 1 - \bar{p} - \bar{q},$$

and $(\bar{p}, \bar{q}) \in (0,1)^2$ such that $1 - \bar{p} - \bar{q} \in (0,1)$. Let $\lambda \in (0,1)$, $\bar{p} = \lambda p$ and $\bar{q} = \lambda(1-p)$ such that $1 - \bar{p} - \bar{q} = 1 - \lambda$. Then,

$$\mathbf{E}\left[s^{\frac{S_t}{S_{t-1}}}\right] = \mathbf{E}\left[s^{\eta_t \Delta N_t + 1}\right] = s^{1+b}\,\bar{p} + s^{1+a}\,\bar{q} + s(1-\lambda) = \mathbf{E}\left[s^{\frac{T_t}{T_{t-1}}}\right],$$

and $S_0 = T_0$. Thus the trinomial and the ternary models are equivalent in law. The introduction of the second one is motivated by the following remark. As explained in Halconruy's PhD dissertation (see [31], conclusion of chapter 4), it turned out to be impossible to derive a Karatzas-Ocone-type hedging formula for replicable claims in the trinomial model (underlying by a sequence of $\{-1,0,1\}$-valued independent variables) by the Clark-Ocone formula stated in Decreusefond and Halconruy (see [20], Theorem 3.3). Indeed, the $\mathcal{F}_k$-measurability of the term $D_k\mathbf{E}\left[F|\mathcal{F}_k\right]$ appearing in this prevents from deriving the expected $\mathcal{F}$-predictable drift process. This observation was prone to replace the trinomial model by a ternary model, based on a jump process, and, as we will see, lends itself more easily to the statement of a hedging formula, directly derived from Clark's one (4.3).

**Remark 5.10** (Incompleteness of the ternary model)**.** As explained in Runggaldier [64], the trinomial tree model is an incomplete market; as expected, so does the ternary model. Indeed, the measure with respect to which the sequence of discounted prices is a $\mathcal{F}$-martingale, is not unique. Considering the process $(S_t)_{t\in\mathbf{T}}$ defined by (5.11) and that is identically distributed to the one of the trinomial model, we expect to reach the same incompleteness result. By writing for any $t \in \{1 : T\}$,

$$\Delta\overline{S}_t = \frac{S_t - (1+r)S_{t-1}}{(1+r)^t} = \frac{\left[b\mathbf{1}_{\{W_t=1\}} + a\mathbf{1}_{\{W_t=-1\}}\right]\Delta N_t - r}{(1+r)^t} \times S_{t-1},$$

it appears that the discounted price sequence is a $\mathcal{F}$-martingale within the condition $\lambda(bp + aq) - r = 0$. As expected, the system

$$\begin{cases} \lambda(bp+aq) &= r \\ p+q &= 1 \end{cases}$$

admits infinitely many solutions $(\lambda, p, q) \in (0,1)^3$ such that any triplet $(\lambda, p, q)$ forms a convex $\mathscr{M}$ set (here a segment) characterized by its extremal points, i.e. the measures $\mathbf{P}^0 = (1, (r - a)/(b-a), (b-r)/(b-a))$ and $= \mathbf{P}^1 = (r/b, 1, 0)$,which are not equivalent to $\mathbf{P}$ but such that any convex combination $\mathbf{P}^\gamma = \gamma\mathbf{P}^0 + (1-\gamma)\mathbf{P}^1$ is. Any measure defined on $\Omega$ and with respect to which the sequence $\overline{\mathrm{S}}$ is a $\mathcal{F}$-martingale is called a *$\mathcal{F}$-martingale measure.*

The value of the portfolio at time $t \in \mathbf{T}$ is given by the random variable

$$\mathrm{V}_t = \alpha_t \, \mathrm{A}_t + \varphi_t \, \mathrm{S}_t,$$

where $(\alpha_t, \varphi_t)_{t\in\mathbf{T}}$ is a couple of $\mathcal{F}$-predictable processes modelling respectively the amounts of riskless and risky assets held in the portfolio.
The aim of this subsection is to exhibit a hedging formula; this is, given a nonnegative $\mathcal{F}_T$-measurable random variable F (called *claim*), to find an *admissible* strategy $\psi = (\alpha, \varphi)$ that is *self-financed* in the sense where for any $t \in \mathbf{T} \setminus \{T\}$,

$$\mathrm{A}_t \left(\alpha_{t+1} - \alpha_t\right) + \mathrm{S}_t \left(\varphi_{t+1} - \varphi_t\right) = 0, \tag{5.12}$$

and which corresponding portfolio value satisfies $\mathrm{V}_0 > 0$, $\mathrm{V}_t \geqslant 0$ for all $t \in \mathbf{T} \setminus \{T\}$, and $\mathrm{V}_T = \mathrm{F}$. In an incomplete market, there is no systematic hedging formula, since all claims are not reachable; they have an *intrisic risk*. Face to the impossibility to perform a perfect hedge in the general case, we can only hope to reduce the a priori risk to this minimal component. The question of hedging in an incomplete market has been widely investigated for years (see for instance Dalang [16], Föllmer and Schweizer [26] in continuous time, Schweizer [70] in discrete time). As explained in Remark 5.9 the ternary model is not complete; we choose to deal with the optimization problem in return:

$$\min_{\psi\in\mathscr{S}} \mathbf{E}\left[(\mathrm{F} - x - \mathrm{V}_T(\psi))^2\right], \tag{5.13}$$

where the claim F and the initial capital $x \in \mathbf{R}_+^*$ are given, and $\mathscr{S}$ is the set of $\mathcal{F}$-predictable admissible strategies. The *mean-variance tradeoff* process $(\mathrm{K}_t)_{t\in\mathbf{T}}$ is defined by

$$\mathrm{K}_t = \sum_{s=1}^{t} \frac{\left(\mathbf{E}\left[\Delta\mathrm{S}_s \,|\, \mathcal{F}_{s-1}\right]\right)^2}{\mathrm{var}[\Delta\mathrm{S}_s \,|\, \mathcal{F}_{s-1}]} \; ; \; t \in \mathbf{T}.$$

Introduce also the discrete analogue of the *minimal martingale measure* (see Föllmer and Schweizer [26]), i.e. the signed measure $\widehat{\mathbf{P}}$ defined on $(\Omega, \mathcal{F})$ such that

$$\frac{\mathrm{d}\widehat{\mathbf{P}}}{\mathrm{d}\mathbf{P}} = \prod_{t=1}^{\mathbf{T}} \frac{1 - \theta_t \Delta\mathrm{S}_t}{1 - \theta_t \mathbf{E}\left[\Delta\mathrm{S}_t | \mathcal{F}_{t-1}\right]}, \tag{5.14}$$

where $(\theta_t)_{t\in\mathbf{T}^*}$ is the $\mathcal{F}$-predictable process such that $\theta_t = \mathbf{E}\left[\Delta\mathrm{S}_t \,|\, \mathcal{F}_{t-1}\right] / \mathbf{E}\left[(\Delta\mathrm{S}_t)^2 \,|\, \mathcal{F}_{t-1}\right]$, for any $t \in \mathbf{T}^*$. Last, consider the Kunita-Watanabe decomposition of F (see Metivier [41] or Schweizer [70]) i.e. the unique couple of processes $(\xi^{\mathrm{F}}, \mathrm{L}^{\mathrm{F}})$ where $\xi^{\mathrm{F}}$ is a square-integrable admissible strategy and $\mathrm{L}^{\mathrm{F}}$ is a $\mathcal{F}$-martingale, strongly orthogonal to S, with null initial value and such that

$$\mathrm{F} = \mathrm{F}_0 + \sum_{t\in\mathbf{T}} \xi_t^{\mathrm{F}} \Delta \mathrm{S}_t + \mathrm{L}_T^{\mathrm{F}} \quad \mathbf{P}\text{-a.s.}$$

Within previous notations, M. Schweizer gives in ([69], Proposition 4.3) an expression of the quadratic-loss minimizing strategy.

**Theorem 5.11** (Schweizer, 1992)**.** *Provided* $(\mathrm{K}_t)_{t\in\mathbf{T}}$ *is deterministic, the solution of* (5.13) *is given by*

$$\varphi_t^* = \xi_t^{\mathrm{F}} + \frac{\mathbf{E}\left[\Delta \mathrm{S}_t \,|\, \mathcal{F}_{t-1}\right]}{\mathbf{E}\left[(\Delta \mathrm{S}_t)^2 \,|\, \mathcal{F}_{t-1}\right]} (\widehat{\mathbf{E}}\left[\mathrm{F}|\mathcal{F}_t\right] - x - \mathrm{V}_{t-1}(\varphi^*)) \tag{5.15}$$

*where* $\widehat{\mathbf{E}}$ *denotes the expectation with respect to the measure* $\widehat{\mathbf{P}}$*. i.e. the minimal martingale measure defined by* (5.14)*. Moreover, the quota of the riskless asset* $(\mathrm{A}_t)_{t\in\mathbf{T}}$ *is given by* $\alpha_0 = \widehat{\mathbf{E}}\left[\mathrm{F}\right]/\mathrm{S}_0$ *and for any* $t \in \mathbf{T}^*$*,*

$$\alpha_t = \alpha_{t-1} - \frac{(\varphi_t - \varphi_{t-1})\mathrm{S}_{t-1}}{\mathrm{A}_{t-1}}.$$

**Remark 5.12.** If the contingent claim F is reachable, then $\varphi^* = \xi^{\mathrm{F}}$. The term $\xi^{\mathrm{F}}$ in (5.15) can be interpreted as a pure hedging demand, whereas the second one can be viewed as a demand for mean-variance purposes (see Schweizer [69]).

These results are slot to our formalism to solve (5.13) in the ternary model.

**Lemma 5.13.** *The mean-variance tradeoff process of the ternary model is deterministic.*

*Proof.* For any $t \in \mathbf{T}$,

$$\frac{\left(\mathbf{E}\left[\Delta \mathrm{S}_t \,|\, \mathcal{F}_{t-1}\right]\right)^2}{\mathrm{var}[\Delta \mathrm{S}_t \,|\, \mathcal{F}_{t-1}]} = \frac{\left(\mathbf{E}\left[\eta_t \Delta \mathrm{N}_t|\mathcal{F}_{t-1}\right]\right)^2}{\mathrm{var}[\eta_t \Delta \mathrm{N}_t|\mathcal{F}_{t-1}]} = \frac{\lambda(bp+aq)^2}{p(1-\lambda p)b^2 - 2ab\lambda pq + a^2 q(1-\lambda q)},$$

is a deterministic constant. □

The family $\mathcal{R}$ is provided by Gram-Schmidt process (2.17) such that

$$\Delta \mathrm{R}_{(t,1)} = \Delta \mathrm{Z}_{(t,1)} \quad \text{and} \quad \Delta \mathrm{R}_{(t,-1)} = \Delta \mathrm{Z}_{(t,-1)} + \frac{\lambda^2 pq}{\lambda p(1-\lambda p)} \Delta \mathrm{R}_{(t,1)} = \Delta \mathrm{Z}_{(t,-1)} + \rho \Delta \mathrm{Z}_{(t,1)},$$

where $\rho := \lambda pq/(p(1-\lambda p))$.

**Lemma 5.14** (Kunita-Watanabe decomposition in the ternary model)**.** *For any claim* $\mathrm{F} \in \mathrm{L}^2(\mathbf{P})$ *there exist a square-integrable admissible strategy* $\xi^{\mathrm{F}}$ *and a* $\mathcal{F}$*-martingale* $\mathrm{L}^{\mathrm{F}}$*, strongly orthogonal to* S*, with null intial value such that*

$$\mathrm{F} = \mathrm{F}_0 + \sum_{t\in\mathbf{T}} \xi_t^{\mathrm{F}} \Delta \mathrm{S}_t + \mathrm{L}_T^{\mathrm{F}} \quad \mathbf{P}\text{-}a.s.$$

*Moreover, for any* $t \in \mathbf{T}^*$*,*

$$\xi_t^{\mathrm{F}} = \frac{1}{\mathrm{S}_{t-1} v_t} \Big( \sum_{k\in\mathrm{E}} u_t \widehat{\mathbf{E}}\left[\mathrm{D}_{(t,k)}\mathrm{F}|\mathcal{F}_{t-1}\right] \Big) \quad \text{and} \quad \mathrm{L}_t^{\mathrm{F}} = \mathbf{E}\Big[\mathrm{F} - \sum_{s\in\mathbf{T}} \xi_s^{\mathrm{F}} \Delta \mathrm{S}_s \,\Big|\, \mathcal{F}_t\Big] - \mathbf{E}\Big[\mathrm{F} - \sum_{s\in\mathbf{T}} \xi_s^{\mathrm{F}} \Delta \mathrm{S}_s\Big], \tag{5.16}$$

*where* $\mathbf{E}\left[\mathrm{L}_0^{\mathrm{F}}\right]=0$, *the sequences* $u=(u_t)_{t\in\mathbf{T}}$ *and* $v=(v_t)_{t\in\mathbf{T}}$ *are defined by*

$$u_t=\lambda\big[(b-a\rho)p(1-\lambda p)+a\big(q(1-\lambda q)-\rho\lambda pq\big)\big] \quad \textit{and} \quad v_t=\lambda(b^2p+a^2q).$$

*The minimal martingale measure* $\widehat{\mathbf{P}}$, *equivalent to* $\mathbf{P}$ *can be explicitly given by*

$$\frac{\mathrm{d}\widehat{\mathbf{P}}}{\mathrm{d}\mathbf{P}}=\prod_{t\in\mathbf{T}}\frac{1-\theta_t\Delta\mathrm{S}_t}{1-\theta_t\mathbf{E}\left[\Delta\mathrm{S}_t|\mathcal{F}_{t-1}\right]}, \tag{5.17}$$

*with*

$$\theta_t=\frac{\mathrm{S}_{t-1}(1+\lambda(bp+aq))}{\mathrm{S}_{t-1}^2(1+\lambda(b^2p+a^2q))}=\frac{1+\lambda(bp+aq)}{\mathrm{S}_{t-1}(1+\lambda(b^2p+a^2q))}\ ;\ t\in\mathbf{T}.$$

**Remark 5.15.** The expression (5.16) of $\xi^{\mathrm{F}}$ which is the replicating strategy when F is reachable, is not so dissimilar to that of the hedging strategy in the binomial model (see Privault [56], proposition 1.14.4).

**Theorem 5.16** (Loss quadratic minimizing strategy in the ternary model)**.** *Let* $\widehat{\mathbf{P}}$ *be the minimal martingale measure defined by* (5.17) *and let a claim* F*. The quadratic loss minimizing hedge* $\varphi^*$ *is given by*

$$\varphi_t^*=\xi_t^{\mathrm{F}}+\theta_t\big(\widehat{\mathbf{E}}\big[\mathrm{F}|\mathcal{F}_t\big]-x-\mathrm{V}_{t-1}(\varphi^*)\big),$$

*where* $\xi^{\mathrm{F}}\in\mathscr{S}$ *is given by the Kunita-Watanabe decomposition.*

*Proof.* Since the mean-variance process is deterministic by Lemma 5.13, it suffices to incorporate the result of Lemma 5.14 to Theorem 5.11. The process $(\alpha_t)_{t\in\mathbf{T}}$ is defined by the self-financing condition (5.12). □

# 6 Proofs

## 6.1 Proofs of the section 2

*Proof of Proposition 2.4.* Let $u,v\in\mathcal{P}$; there exists $T\in\mathbf{N}$ such that $u$ and $v$ are of the form (2.5). For any $t\in\{1:T\}$,

$$\begin{aligned}
&\mathbf{E}\Big[\mathrm{J}_1(u\mathbf{1}_{[t,\infty)}\,;\mathcal{R})\mathrm{J}_n(v\mathbf{1}_{[t,\infty)}\,;\mathcal{R})\,\Big|\,\mathcal{F}_{t-1}\Big]\\
&\quad=\sum_{(s,k)\in\{t:T\}\times\mathrm{E}}\ \sum_{(r,\ell)\in\{t:T\}\times\mathrm{E}}\mathbf{E}\left[\mathfrak{u}(\eta,(s,k))\mathfrak{v}(\eta,(r,\ell))\Delta\mathrm{R}_{(s,k)}\Delta\mathrm{R}_{(r,\ell)}\,\middle|\,\mathcal{F}_{t-1}\right]\\
&\quad=\sum_{(s,k,\ell)\in\{t:T\}\times\mathrm{E}^2}\mathbf{E}\left[\mathfrak{u}(\eta,(s,k))\mathfrak{v}(\eta,(s,\ell))\mathbf{E}\left[\Delta\mathrm{R}_{(s,k)}\Delta\mathrm{R}_{(s,\ell)}\,|\,\mathcal{F}_{s-1}\right]\,\middle|\,\mathcal{F}_{t-1}\right]\\
&\qquad+\sum_{(s,k)\in\{t:T\}\times\mathrm{E}}\ \sum_{\substack{(r,\ell)\in\{t:T\}\times\mathrm{E}\\ r>s}}\mathbf{E}\left[\mathfrak{u}(\eta,(s,k))\mathfrak{v}(\eta,(r,\ell))\Delta\mathrm{R}_{(s,k)}\mathbf{E}\left[\Delta\mathrm{R}_{(r,\ell)}|\mathcal{F}_{r-1}\right]\,\middle|\,\mathcal{F}_{t-1}\right]\\
&\qquad+\sum_{(r,\ell)\in\{t:T\}\times\mathrm{E}}\ \sum_{\substack{(s,k)\in\{t:T\}\times\mathrm{E}\\ s>r}}\mathbf{E}\left[\mathfrak{u}(\eta,(s,k))\mathfrak{v}(\eta,(r,\ell))\Delta\mathrm{R}_{(r,\ell)}\mathbf{E}\left[\Delta\mathrm{R}_{(s,k)}|\mathcal{F}_{s-1}\right]\,\middle|\,\mathcal{F}_{t-1}\right]\\
&\quad=\sum_{(s,k)\in\{t:T\}\times\mathrm{E}}\kappa_k\mathbf{E}\left[\mathfrak{u}(\eta,(s,k))\mathfrak{v}(\eta,(s,k))\,\middle|\,\mathcal{F}_{t-1}\right]=\mathbf{E}\left[\langle u\mathbf{1}_{[t,\infty)},v\mathbf{1}_{[t,\infty)}\rangle_{\mathrm{L}^2(\mathbb{X},\tilde{\nu})}\,\middle|\,\mathcal{F}_{t-1}\right],
\end{aligned}$$

where the second and third sums in the second equality vanish as $\mathbf{E}\left[\Delta \mathrm{R}_{(t,k)}\right] = 0$ for all $(t,k) \in \mathbb{X}_T$. The extension of the stochastic integral to the set of square-integrable adapted processes comes from a Cauchy sequence argument. Define the sequence $(u^n)_{n\in\mathbf{N}}$ of simple predictable processes by

$$u^n(\eta,(t,k)) = u(\eta,(t,k))\mathbf{1}_{\{t\in\{1:n\}\}}\mathbf{1}_{\{|u(\eta,(t,k))|\leqslant n\}}.$$

Thus, $(\mathrm{J}_1(u^n\,;\mathcal{R})_{n\in\mathbf{N}}$ is Cauchy and converges in $\mathrm{L}^2(\mathbf{P})$. Let then

$$\mathrm{J}_1(u\,;\mathcal{R}) = \lim_{n\to\infty} \mathrm{J}_1(u^n\,;\mathcal{R}).$$

The limit is independent of the approximating sequence by applying the isometry property (2.6) with $t = 1$. Hence the result. □

*Proof of Proposition 2.6.* Assume with no loss of generality that $m > n$. Let $(f_n, g_m) \in \mathrm{L}^2(\mathbb{X},\tilde{\nu})^{\circ n} \times \mathrm{L}^2(\mathbb{X},\tilde{\nu})^{\circ m}$. For any $(\mathbf{t}_n,\mathbf{s}_m) \in \mathbf{N}^{n,<} \times \mathbf{N}^{m,<}$, there exists $j_0 \in \mathbf{N}$ such that $s_{j_0} \in \mathbf{s}_m \setminus \mathbf{t}_n$. By independence of the random variable $\Delta\mathrm{R}_{(s_{j_0},k_{j_0})}$ with respect to the $\sigma$-algebra $\mathcal{G}_{s_{j_0}} = \sigma\big\{\sum_{(s,k)} \eta(s,k),\ s \neq s_{j_0},\ k \in \mathrm{E}\big\}$,

$$\begin{aligned}
\mathbf{E}\left[\mathrm{J}_n(f_n\,;\mathcal{R})\,\mathrm{J}_m(g_m\,;\mathcal{R})\right] &= n!\,m! \sum_{(\mathbf{t}_n,\mathbf{k}_n)\in\mathbb{X}^{n,<}} \sum_{(\mathbf{s}_m,\mathbf{l}_m)\in\mathbb{X}^{m,<}} f_n(\mathbf{t}_n,\mathbf{k}_n)\, g_m(\mathbf{s}_m,\mathbf{l}_m) \\
&\qquad \times \mathbf{E}\left[\prod_{i=1}^n\prod_{j=1}^m \Delta\mathrm{R}_{(t_i,k_i)}\Delta\mathrm{R}_{(s_j,\ell_j)}\right] \\
&= n!\,m! \sum_{(\mathbf{t}_n,\mathbf{k}_n)\in\mathbb{X}^{n,<}} \sum_{(\mathbf{s}_m,\mathbf{l}_m)\in\mathbb{X}^{m,<}} f_n(\mathbf{t}_n,\mathbf{k}_n)\, g_m(\mathbf{s}_m,\mathbf{l}_m) \\
&\qquad \times \mathbf{E}\left[\Delta\mathrm{R}_{(s_{j_0},k_{j_0})}\right] \prod_{i=1}^n \prod_{\substack{j=1\\ j\neq j_0}}^m \mathbf{E}\left[\Delta\mathrm{R}_{(t_i,k_i)}\Delta\mathrm{R}_{(s_j,\ell_j)}\right] = 0,
\end{aligned}$$

and for $m = n$,

$$\begin{aligned}
\mathbf{E}\left[\mathrm{J}_n(f_n\,;\mathcal{R})\mathrm{J}_n(g_n\,;\mathcal{R})\right] &= (n!)^2 \sum_{(\mathbf{t}_n,\mathbf{k}_n)\in\mathbb{X}^{n,<}} \sum_{(\mathbf{s}_n,\mathbf{l}_n)\in\mathbb{X}^{n,<}} f_n(\mathbf{t}_n,\mathbf{k}_n)\, g_n(\mathbf{s}_n,\mathbf{l}_n) \\
&\qquad \times \mathbf{E}\left[\prod_{i,j=1}^n \Delta\mathrm{R}_{(t_i,k_i)}\Delta\mathrm{R}_{(s_j,\ell_j)}\right] \\
&= (n!)^2 \sum_{\substack{(\mathbf{t}_n,\mathbf{k}_n)\in\mathbb{X}^{n,<}\\ \mathbf{l}_n\in\mathrm{E}^{n,<}}} f_n(\mathbf{t}_n,\mathbf{k}_n)\, g_n(\mathbf{t}_n,\mathbf{l}_n)\mathbf{E}\left[\prod_{i,j=1}^n \Delta\mathrm{R}_{(t_i,k_i)}\Delta\mathrm{R}_{(t_i,\ell_i)}\right] \\
&= n!\,\langle f_n, g_n\rangle_{\mathrm{L}^2(\mathbb{X},\tilde{\nu})^{\circ n}},
\end{aligned}$$

since $\mathbf{E}\left[\Delta\mathrm{R}_{(t,k)}\,\Delta\mathrm{R}_{(t,\ell)}\right] = \kappa_k \mathbf{1}_{\{k\}}(\ell)$. Besides, for any $f_n \in \mathrm{L}^2(\mathbb{X},\tilde{\nu})^{\circ n}$,

$$\mathrm{J}_n(f_n\,;\mathcal{R}) = n \sum_{(t,k)\in\mathbb{X}} \mathrm{J}_{n-1}\,(\pi^n_{(t,k)} f_n;\mathcal{R})\,\Delta\mathrm{R}_{(t,k)}$$

$$= n! \sum_{(t,k)\in\mathbb{X}} \sum_{(\mathbf{t}_{n-1},\mathbf{k}_{n-1})\in\mathbb{X}^{n,<}} f_n\big((\mathbf{t}_{n-1},\mathbf{k}_{n-1}),(t,k)\big)\, \Delta \mathrm{R}_{(t,k)} \prod_{i=1}^{n-1} \Delta \mathrm{R}_{(t_i,k_i)}$$

$$= n! \sum_{(\mathbf{t}_n,\mathbf{k}_n)\in\mathbb{X}^{n,<}} f_n(\mathbf{t}_n,\mathbf{k}_n) \prod_{i=1}^{n} \Delta \mathrm{R}_{(t_i,k_i)},$$

that completes the proof. □

*Proof of Lemma 2.7.* Let $g \in \mathrm{L}^2(\mathbb{X}_{T'})$ and $f_n \in \mathrm{L}^2(\mathbb{X}_{T''})^{\circ n}$ for some $T', T'' \in \mathbf{N}$, the definition of the symmetric tensor product implies $g \circ f_n \in \mathrm{L}^2(\mathbb{X}_T)^{\circ n+1}$ where $T := \max(T', T'')$, and

$$\begin{aligned}
\mathrm{J}_{n+1}(g\circ f_n\,;\mathcal{R}) &= n! \sum_{i=1}^{n+1} \sum_{(\mathbf{t}_{n+1},\mathbf{k}_{n+1})\in\mathbb{X}_T^{n+1,<}} g(t_i,k_i) f_n^{\neg i}(\mathbf{t}_{n+1},\mathbf{k}_{n+1}) \prod_{i=1}^{n+1} \Delta \mathrm{R}_{(t_i,k_i)} \\
&= n! \sum_{i=1}^{n} \sum_{(\mathbf{t}_{n+1},\mathbf{k}_{n+1})\in\mathbb{X}_T^{n+1,<}} g(t_i,k_i) f_n^{\neg i}(\mathbf{t}_{n+1},\mathbf{k}_{n+1}) \prod_{i=1}^{n+1} \Delta \mathrm{R}_{(t_i,k_i)} \\
&\quad + n! \sum_{(t,k)\in\mathbb{X}_T} \sum_{\substack{(\mathbf{t}_n,\mathbf{k}_n)\in\mathbb{X}_T^{n,<}: \\ (t,k)\notin(\mathbf{t}_n,\mathbf{k}_n)}} g(t,k) f_n^{\neg i}(\mathbf{t}_n,\mathbf{k}_n)\, \Delta \mathrm{R}_{(t,k)} \prod_{i=1}^{n} \Delta \mathrm{R}_{(t_i,k_i)} \\
&= n! \sum_{(t,k)\in\mathbb{X}_T} \sum_{i=1}^{n} \sum_{\substack{(\mathbf{t}_n,\mathbf{k}_n)\in\mathbb{X}_T^{n,<}: \\ (t,k)\notin(\mathbf{t}_n,\mathbf{k}_n)}} g(t_i,k_i) f_n^{\neg i}((\mathbf{t}_n,\mathbf{k}_n),(s,k)) \Delta \mathrm{R}_{(s,k)} \prod_{i=1}^{n} \Delta \mathrm{R}_{(t_i,k_i)} \\
&\quad + \sum_{(t,k)\in\mathbb{X}_T} g(t,k) \mathrm{J}_n(f_n \mathbf{1}_{\{1:t-1\}^n}\,;\mathcal{R}) \Delta \mathrm{R}_{(t,k)} \\
&= n \sum_{(t,k)\in\mathbb{X}_T} \mathrm{J}_n\big(f_n(\star,(t,k)) \circ g(\cdot) \mathbf{1}_{\{1:t-1\}^n}(\star,\cdot)\,;\mathcal{R}\big) \Delta \mathrm{R}_{(t,k)} \\
&\quad + \sum_{(t,k)\in\mathbb{X}_T} g(t,k) \mathrm{J}_n(f_n \mathbf{1}_{\{1:t-1\}^n}\,;\mathcal{R}) \Delta \mathrm{R}_{(t,k)}.
\end{aligned}$$

The result is then extended to $g \in \mathrm{L}^2(\mathbb{X})$ and $f_n \in \mathrm{L}^2(\mathbb{X})^{\circ n}$ by a standard Cauchy argument. □

*Proof of Lemma 2.8.* Let $T \in \mathbf{N}$ and $f_n \in \mathrm{L}^2(\mathbb{X}_T)^{\circ n}$. For any $t \in \mathbf{N}$ such that $t < T$,

$$\begin{aligned}
\mathbf{E}\left[\mathrm{J}_n(f_n\,;\mathcal{R}) \,|\, \mathcal{F}_t\right] &= n! \sum_{(\mathbf{t}_n,\mathbf{k}_n)\in(\mathbb{X}_T)^{n,<}} f_n(\mathbf{t}_n,\mathbf{k}_n)\, \mathbf{E}\left[\prod_{i=1}^{n} \Delta \mathrm{R}_{(t_i,k_i)} \,\middle|\, \mathcal{F}_t\right] \\
&= n! \sum_{(\mathbf{t}_n,\mathbf{k}_n)\in(\mathbb{X}_t)^{n,<}} f_n(\mathbf{t}_n,\mathbf{k}_n)\, \mathbf{E}\left[\prod_{i=1}^{n} \Delta \mathrm{R}_{(t_i,k_i)} \,\middle|\, \mathcal{F}_t\right] = \mathrm{J}_n\big(f_n \mathbf{1}_{\{1:t\}}\,;\mathcal{R}\big),
\end{aligned}$$

since the independence of the centered variables $\{\Delta \mathrm{R}_{(t_i,k_i)},\ (t_i,k_i) \in \mathbb{X},\ i \in \{1:n\}\}$ implies that $\mathbf{E}\big[\prod_{i=1}^{n} \Delta \mathrm{R}_{(t_i,k_i)} \big| \mathcal{F}_t\big] = 0$ if there exists $i_0 \in \{1:T\}$ such that $t_{i_0} > t$. The result is extended to $\mathrm{L}^2(\mathbb{X})^{\circ n}$ by a limit procedure. □

*Proof of Lemma 2.10.* It suffices to note that $\mathcal{H}_s \cap \mathcal{L}^0(\mathbf{P}, \mathcal{F}_t)$ for some $(s,t) \in \mathbf{N}$, $s \leqslant t$, is generated by the orthogonal family

$$\{1\} \cup \left\{ \prod_{i=1}^{s} \Delta \mathrm{R}_{(t_i,k_i)},\ 1 \leqslant t_1 < \cdots < t_s \leqslant t,\ (k_1, \dots, k_s) \in \mathrm{E}^s \right\}. \tag{6.1}$$

Indeed any element of $\overline{\mathcal{R}}_t = \mathrm{Span}\{\Delta \mathrm{R}_{(s,k)},\ (s,k) \in \mathbb{X}_t\}$ can be expressed in terms of multiple integrals as

$$\prod_{i=1}^{s} \Delta \mathrm{R}_{(t_i,k_i)} = \mathrm{J}_s\Big(\mathbf{1}^{<}_{\{(t_1,k_1),\dots,(t_s,k_s)\}} \mathbf{1}_{\{0:t\}^s}\Big).$$

We conclude by noting that the dimensions of $\overline{\mathcal{R}}_t$ and $\mathcal{L}^0(\mathbf{P}, \mathcal{F}_t)$ in (2.12) are both equal to

$$1 + \sum_{s=1}^{t} |\mathrm{E}|^s \times \binom{t}{s} = (1 + |\mathrm{E}|)^t.$$

The proof is thus complete. □

*Proof of Proposition 2.16.* Let, for notation purposes, $\mathfrak{m}^{-1}_{k^i,k^j}$ designate the $(i,j)$-th entry of matrix $\mathfrak{M}^{-1}$, that is the inverse of matrix $\mathfrak{M}$ defined by (2.18). It suffices to state it for any random variable $\mathrm{F} \in \mathcal{S}$. Let $\mathrm{E} = \{k^1, \dots, k^{\overline{\mathfrak{m}}}\}$. By Theorem (2.11), The chaotic decomposition of F reads

$$\mathrm{F} = \mathbf{E}\left[\mathrm{F}\right] + \sum_{\mathbf{t}_n} \sum_{i_1=1}^{\overline{\mathfrak{m}}} \cdots \sum_{i_n=1}^{\overline{\mathfrak{m}}} f_n((t_1, k_1^{i_1}), \dots, (t_n, k_n^{i_n})) \prod_{j=1}^{n} \Delta \mathrm{R}_{(t_j, k_j^{i_j})}.$$

Since $\Delta \mathrm{R}_{(t_j,k_j^\ell)} = \sum_{p=1}^{\ell} \mathfrak{m}^{-1}_{k^\ell k^p}\, \Delta \mathrm{Z}_{(t_j,k_j^p)}$, we get

$$\begin{aligned}
\mathrm{F} - \mathbf{E}\left[\mathrm{F}\right] &= \sum_{\mathbf{t}_n} \sum_{i_1=1}^{\overline{\mathfrak{m}}} \cdots \sum_{i_n=1}^{\overline{\mathfrak{m}}} f_n((t_1, k_1^{i_1}), \dots, (t_n, k_n^{i_n})) \prod_{j=1}^{n} \Big( \sum_{p=1}^{i_j} \mathfrak{m}^{-1}_{k^{i_j} k^p}\, \Delta \mathrm{Z}_{(t_j,k_j^p)} \Big) \\
&= \sum_{\mathbf{t}_n} \sum_{i_1=1}^{\overline{\mathfrak{m}}} \cdots \sum_{i_n=1}^{\overline{\mathfrak{m}}} f_n((t_1, k_1^{i_1}), \dots, (t_n, k_n^{i_n})) \Big( \sum_{p_1=1}^{i_1} \cdots \sum_{p_n=1}^{i_n} \prod_{j=1}^{n} \mathfrak{m}^{-1}_{k^{i_j} k^{p_j}}\, \Delta \mathrm{Z}_{(t_j,k_j^{p_j})} \Big) \\
&= \sum_{\mathbf{t}_n} \sum_{p_1=1}^{\overline{\mathfrak{m}}} \cdots \sum_{p_n=1}^{\overline{\mathfrak{m}}} \sum_{i_1=p_1}^{\overline{\mathfrak{m}}} \cdots \sum_{i_n=p_n}^{\overline{\mathfrak{m}}} f_n((t_1, k_1^{i_1}), \dots, (t_n, k_n^{i_n})) \prod_{j=1}^{n} \Big( \mathfrak{m}^{-1}_{k^{i_j} k^{p_j}}\, \Delta \mathrm{Z}_{(t_j,k_j^{p_j})} \Big) \\
&= \sum_{\mathbf{t}_n} \sum_{p_1=1}^{\overline{\mathfrak{m}}} \cdots \sum_{p_n=1}^{\overline{\mathfrak{m}}} \Big( \sum_{i_1=p_1}^{\overline{\mathfrak{m}}} \cdots \sum_{i_n=p_n}^{\overline{\mathfrak{m}}} \prod_{j=1}^{n} \mathfrak{m}^{-1}_{k^{i_j} k^{p_j}}\, f_n((t_1, k_1^{i_1}), \dots, (t_n, k_n^{i_n})) \Big) \prod_{j=1}^{n} \Delta \mathrm{Z}_{(t_j,k_j^{p_j})},
\end{aligned}$$

where we summed over the set of $\{\mathbf{t}_n \in \mathbb{X}_T^{n,<} \ :\ \mathbf{t}_n = (t_1, \cdots, t_n)\}$. The result is extended to $\mathrm{L}^2(\mathbf{P})$ by density of $\mathcal{S}$. □

*Proof of Proposition 2.17.* For any $T \in \mathbf{N}, t \in \{1 : T\}$ define

$$\zeta_t^T = 1 + \sum_{n=1}^{T} \frac{1}{n!} \mathrm{J}_n(h^{\otimes n} \mathbf{1}_{\{1:t\}^n}).$$

where we assume with no loss of generality that $\mathbf{E}\left[\zeta_t^T\right] = 1$. Consider $T$ large enough such that $T > t$; then

$$
\begin{aligned}
1 &+ \sum_{s=1}^{t}\sum_{k\in \mathrm{E}} h(s,k)\zeta_{s-1}^{T}\Delta \mathrm{R}_{(s,k)} \\
&= 1 + \sum_{s=1}^{t}\sum_{k\in \mathrm{E}} h(s,k)\Big(1 + \sum_{n=1}^{T}\frac{1}{n!}\mathrm{J}_n\big(h^{\otimes n}\mathbf{1}_{\{1:s-1\}^n}\big)\Big)\Delta \mathrm{R}_{(s,k)} \\
&= 1 + \sum_{s=1}^{t}\sum_{k\in \mathrm{E}} h(s,k)\Delta \mathrm{R}_{(s,k)} + \sum_{n=1}^{T}\frac{1}{n!}\mathrm{J}_{n+1}\big(h^{\otimes n+1}\mathbf{1}_{\{1:t\}^{n+1}}\big) \\
&\quad - \sum_{n=1}^{T}\sum_{k\in \mathrm{E}}\frac{n}{n!}\sum_{s=1}^{t}\mathrm{J}_n\big(h^{\otimes n}(\star,(s,k))\circ h(\cdot)\mathbf{1}_{\{1:s-1\}^n}(\star,\cdot)\big)\Delta \mathrm{R}_{(s,k)} \\
&= 1 + \mathrm{J}_1(h\mathbf{1}_{\{1:t\}}) + \sum_{n=1}^{T}\frac{1}{n!}\mathrm{J}_{n+1}\big(h^{\otimes n+1}\mathbf{1}_{\{1:t\}^{n+1}}\big) - \sum_{n=1}^{T}\sum_{k\in \mathrm{E}}\frac{n}{n!}\mathrm{J}_n\big(h^{\otimes n+1}(\star,(s,k))\mathbf{1}_{\{1:t\}^{n+1}}(\star)\big)\Delta \mathrm{R}_{(s,k)} \\
&= 1 + \mathrm{J}_1(h\mathbf{1}_{\{1:t\}}) + \sum_{n=1}^{T}\frac{1}{n!}\mathrm{J}_{n+1}\big(h^{\otimes n+1}\mathbf{1}_{\{1:t\}^{n+1}}\big) - \sum_{n=1}^{T}\frac{n}{(n+1)!}\mathrm{J}_{n+1}(h^{\otimes n+1}\mathbf{1}_{\{1:t\}^{n+1}}) \\
&= 1 + \mathrm{J}_1(h\mathbf{1}_{\{1:t\}}) + \sum_{n=2}^{T+1}\frac{1}{n!}\mathrm{J}_n\big(h^{\otimes n}\mathbf{1}_{\{1:t\}^n}\big) = \zeta_t^{T+1},
\end{aligned}
$$

where we used Lemma 2.7 in the second line and the definition of the multiple integral (2.7) in the penultimate one. Since by the very definition of Doléans exponential (2.20), for all $t \in \mathbf{N}$ $\zeta_t^T$ tends to $\xi_t(h)$ almost surely when $T$ goes ton infinity, we get

$$\xi_t(h) = 1 + \sum_{s=1}^{t}\Big(\sum_{k\in \mathrm{E}} h(s,k)\Delta \mathrm{R}_{(s,k)}\Big)\xi_{s-1}(h).$$

Besides, the sequence $(\zeta_t)_{t\in\mathbf{N}}$ satisfies the equation in differences

$$\xi_t(h) - \xi_{t-1}(h) = \xi_{t-1}(h)\sum_{k\in \mathrm{E}} g(t,k)\big(\mathbf{1}_{(t,k)} - \lambda\mathbf{Q}(\{k\})\big)$$

where $\mathrm{J}_1(h) = \mathrm{J}_1(g\,;\mathcal{Z})$. On the other hand, provided the product converges, define the sequence of exponential products $(\xi_t^{\mathcal{Z}}(g))_{t\in\mathbf{N}}$, that stand for the Doléan exponentials with respect to the family $\mathcal{Z}$, by

$$\xi_t(h) = \xi_t^{\mathcal{Z}}(g) = \prod_{t\in\mathbf{N}}\Big(1 + \sum_{k\in \mathrm{E}} g(t,k)\big(\mathbf{1}_{(t,k)} - \lambda\mathbf{Q}(\{k\})\big)\Big)$$

and so that for all $t \in \mathbf{N}$,

$$\xi_t^{\mathcal{Z}}(g) = 1 + \sum_{s=1}^{t}\Big(\sum_{k\in \mathrm{E}} h(s,k)\Delta \mathrm{R}_{(s,k)}\Big)\xi_{s-1}^{\mathcal{Z}}(g).$$

By uniqueness of the decomposition, provided the series and product converge, $\xi_t^{\mathcal{Z}}(g) = \xi_t(h)$ for any $t \in \mathbf{N}$; that leads to the conclusion. □

## 6.2 Proofs of Section 3

### 6.2.1 Proofs of Subsection 3.1

*Proof of Lemma 3.1.* Let $\mathbf{K}$ be a probability measure on $\mathbf{N}$, $\mathrm{V}_1, \mathrm{V}_2, \ldots$ independent random elements in E with distribution $\mathbf{Q}$, and K a random variable with distribution $\mathbf{K}$ supposed to be independent of $(\mathrm{V}_n,\, n \in \mathbf{N})$. Recall that

$$\varpi = \sum_{j=1}^{\mathrm{K}} \delta_{\mathrm{V}_j} \tag{6.2}$$

is called a mixed binomial process with mixing distribution $\mathbf{K}$ and sampling distribution $\mathbf{Q}$. Let $\eta \in \widehat{\mathfrak{N}_{\mathbb{X}}}$; there exists $T \in \mathbf{N}$ such that $\eta$ is a marked binomial process on $\mathbb{X}_T$. By its very definition any marked binomial process on $\mathbb{X}_T$ of intensity measure $\nu$ is a mixed binomial process with mixing distribution $\mathcal{B}\mathrm{in}(T, \lambda)$ and sampling distribution $\mathbf{Q}$. Moreover, for any $n \in \{1 : T\}$, $\varpi_{|\mathrm{K}=n}$ is a binomial process of intensity measure $n\mathbf{Q}$. Then, as a special case of the Georgii-Nguyen-Zessin formula (see [17], Proposition 15.5.II with $\mathbf{x} = \varpi_{|\mathrm{K}=n}$ and $\rho = n\nu$), for $\mathfrak{u}$ measurable application from $\mathfrak{N}_{\mathbb{X}} \times \mathbb{X}$ into $[0, +\infty]$,

$$\begin{aligned}
\mathbf{E}\Big[\sum_{(t,k)\in\eta} \mathfrak{u}\big(\eta, (t,k)\big)\Big] &= \sum_{n=1}^{T} \mathbf{E}\left[\sum_{k\in\mathrm{E}} \mathbf{E}\left[\mathfrak{u}(\varpi_{|\mathrm{K}}, k)\, \varpi(k) \,\Big|\, \mathrm{K} = n\right]\right] \\
&= \sum_{n=1}^{T} n \sum_{k\in\mathrm{E}} \mathbf{E}\left[\mathfrak{u}(\varpi_{|\mathrm{K}-1} + \delta_k, k)\right] \mathbf{Q}(\{k\}) \\
&= \sum_{n=1}^{T}\sum_{j=1}^{n}\sum_{k\in\mathrm{E}} \mathbf{E}\Big[\big(\sum_{i\neq j} \delta_{\mathrm{V}_i} + \delta_k, \mathrm{V}_j\big)\Big]\Big|_{\mathrm{V}_j = k} \\
&= \mathbf{E}\left[\int_{\mathbb{X}_T} \mathfrak{u}\big(\pi_t(\eta) + \delta_{(t,k)}, (t,k)\big) \mathrm{d}\nu(t,k)\right],
\end{aligned}$$

where we have used the mixed binomial representation of $\eta$ in the second line and in the last one. Similarly we have

$$\mathbf{E}\Big[\sum_{(t,k)\in\eta} \mathfrak{u}\big(\eta - \delta_{(t,k)}, (t,k)\big)\Big] = \mathbf{E}\left[\int_{\mathbb{X}_T} \mathfrak{u}\big(\pi_t(\eta), (t,k)\big) \mathrm{d}\nu(t,k)\right]$$

Hence the result. □

*Proof of Proposition 3.4.* Under the hypotheses of the proposition, by noting that $\mathrm{D}^{+}_{(t,k)}\mathrm{F} - \overline{\mathrm{D}}_t\mathrm{F} = \mathfrak{f}(\pi_t(\eta) + \delta_{(t,k)}) - \mathfrak{f}(\eta)$, we have

$$\begin{aligned}
&\mathbf{E}\Big[\int_{\mathbb{X}} \Big(\mathrm{D}^{+}_{(t,k)}\mathrm{F} - \overline{\mathrm{D}}_t\mathrm{F}\Big)\, \mathrm{d}\nu(t,k)\Big] \\
&\quad= \mathbf{E}\Big[\int_{\mathbb{X}} [\mathfrak{f}(\pi_t(\eta) + \delta_{(t,k)}) - \mathfrak{f}(\eta)]\, \mathfrak{u}(\eta, (t,k))\, \mathrm{d}\nu(t,k)\Big] \\
&\quad= \mathbf{E}\Big[\int_{\mathbb{X}} \mathfrak{f}(\pi_t(\eta) + \delta_{(t,k)})\, \mathfrak{u}(\eta, (t,k))\, \mathrm{d}\nu(t,k)\Big] - \mathbf{E}\Big[\int_{\mathbb{X}} [\mathfrak{f}(\eta)\, \mathfrak{u}(\eta, (t,k))\, \mathrm{d}\nu(t,k)\Big]
\end{aligned}$$

$$= \mathbf{E}\Big[\int_{\mathbb{X}} \mathfrak{f}(\eta)\, \mathfrak{u}(\pi_t(\eta),(t,k))\, \mathrm{d}\eta(t,k)\Big] - \mathbf{E}\Big[\int_{\mathbb{X}} [\mathfrak{f}(\eta)\, \mathfrak{u}(\eta,(t,k))\, \mathrm{d}\nu(t,k)\Big]$$
$$= \mathbf{E}\Big[\int_{\mathbb{X}} \mathfrak{f}(\eta)\, \mathfrak{u}(\eta,(t,k))\, \mathrm{d}\eta(t,k)\Big] - \mathbf{E}\Big[\int_{\mathbb{X}} [\mathfrak{f}(\eta)\, \mathfrak{u}(\eta,(t,k))\, \mathrm{d}\nu(t,k)\Big] = \mathbf{E}\big[\mathrm{F}\widetilde{\delta}(u)\big],$$

where we have used the Mecke formula in the third line and that $u$ is $\mathcal{F}$-predictable to get the last one. Hence the result. □

### 6.2.2 Proofs of Subsection 3.2

*Proof of Theorem 3.8.* The proof is identical to that of Theorem 1.8.2 in Privault [56]. Let $\mathrm{F} = \mathrm{J}_n(f_n)$ for some $f_n \in \mathrm{L}^2(\mathbb{X})^{\circ n}$ and $u_\cdot = \mathrm{J}_m(g_{m+1}(\star,(\cdot))) \in \mathcal{U}$ for some $g_{m+1} \in \mathrm{L}^2(\mathbb{X})^{\circ m} \otimes \mathrm{L}^2(\mathbb{X})$. Then,

$$\begin{aligned}
&\mathbf{E}\Big[\langle \mathrm{D}_\cdot \mathrm{J}_n(f_n), \mathrm{J}_m(g_{m+1}(\star,\cdot))\rangle_{\mathrm{L}^2(\mathbb{X},\tilde{\nu})}\Big] \\
&= n\mathbf{E}\left[\langle \mathrm{J}_{n-1}(f_n(\star,\cdot))\mathbf{1}_{\Delta_n(\star,\cdot)}, \mathrm{J}_m(g_{m+1}(\star,\cdot))\rangle_{\mathrm{L}^2(\mathbb{X},\tilde{\nu})}\right] \\
&= n!\mathbf{1}_{\{n-1\}}(m) \int_{\mathbb{X}} \mathbf{E}\left[\mathrm{J}_{n-1}(f_n(\star,(t,k)))\mathbf{1}_{\Delta_n(\star,(t,k))}, \mathrm{J}_m(g_{m+1}(\star,(t,k)))\right] \, \mathrm{d}\tilde{\nu}(t,k) \\
&= n!\mathbf{1}_{\{n\}}(m+1)\langle \mathbf{1}_{\Delta_n(\star,(t,k))} f_n(\star,(t,k)), \tilde{g}_{m+1}(\star,(t,k))\rangle_{\mathrm{L}^2(\mathbb{X},\tilde{\nu})} \\
&= \mathbf{E}\left[\mathrm{J}_n(f_n)\mathrm{J}_m(\tilde{g}_{m+1})\right] = \mathbf{E}\left[\mathrm{F}\,\delta u\right].
\end{aligned}$$

Hence the result. □

*Proof of Corollary 3.9.* Let $(\mathrm{F}_n)_{n\in\mathbf{Z}_+}$ be a sequence of random variables defined on $\mathcal{S}$ such that $\mathrm{F}_n$ converges to 0 in $\mathrm{L}^2(\mathbf{P})$ and the sequence $(\mathrm{DF}_n)_{n\in\mathbf{Z}_+}$ converges to $\Lambda$ in $\mathrm{L}^2(\mathbf{P}\otimes\tilde{\nu})$. Let $u$ be a simple process. From the integration by parts formula (3.17),

$$\mathbf{E}\Bigg[\sum_{(t,k)\in\mathbb{X}} \mathrm{D}_{(t,k)}\mathrm{F}_n\ u_{(t,k)}\Bigg] = \mathbf{E}\Bigg[\mathrm{F}_n \sum_{(t,k)\in\mathbb{X}} u_{(t,k)}\Delta\mathrm{R}_{(t,k)}\Bigg],$$

where $\sum_{(t,k)\in\mathbb{X}} u_{(t,k)}\Delta\mathrm{R}_{(t,k)} \in \mathrm{L}^2(\mathbf{P})$. Indeed, the process $(\Delta\mathrm{R}_{(t,k)}u_{(t,k)})_{(t,k)\in\mathbb{X}_T}$ belongs to $\mathrm{L}^2(\Omega\times\mathbb{X},\mathbf{P}\otimes\tilde{\nu})$ since, by the Cauchy-Schwarz inequality,

$$\mathbf{E}\Bigg[\sum_{(t,k)\in\mathbb{X}_T} \big|u_{(t,k)}\Delta\mathrm{R}_{(t,k)}\big|^2\Bigg] \leqslant \sum_{(t,k)\in\mathbb{X}_T} \kappa_k \mathbf{E}\Big[u^2_{(t,k)}\delta_{(t,k)}\Big] < \infty.$$

Then,

$$\langle \Lambda, u\rangle_{\mathrm{L}^2(\mathbf{P}\otimes\tilde{\nu})} = \lim_{n\to\infty} \mathbf{E}\Bigg[\mathrm{F}_n \sum_{(t,k)\in\mathbb{X}} u_{(t,k)}\Delta\mathrm{R}_{(t,k)}\Bigg] = 0,$$

for simple process $u$. It follows that $\Lambda = 0$ and then the operator D is closable from $\mathrm{L}^2(\mathbf{P})$ to $\mathrm{L}^2(\Omega\times\mathbb{X},\mathbf{P}\otimes\tilde{\nu})$. By equivalence of the norms $\|\cdot\|_{\mathrm{L}^2(\mathbb{X},\tilde{\nu})}$ and $\|\cdot\|_{\mathrm{L}^2(\mathbb{X},\nu)}$, this result can be extended to $\mathrm{L}^2(\Omega\times\mathbb{X},\mathbf{P}\otimes\nu)$. □

### 6.2.3 Proofs of Subsection 3.3

*Proof of Proposition 3.11 .* The application of $\mathrm{D}^+$ to $\mathrm{F} = \mathrm{J}_n(f_n) \in \mathcal{S}$, and $(t,k) \in \mathbb{X}_T$, gives

$$\begin{aligned}
\mathrm{D}^+_{(t,k)}\mathrm{J}_n(f_n) &= n! \sum_{(\mathbf{t}_n,\mathbf{k}_n)\in(\mathbb{X}_T)^{n,<}} f_n\big((t_1,k_1),\dots,(t_n,k_n)\big) \prod_{i=1}^{n} \mathrm{D}^+_{(t,k)}\,\Delta\mathrm{R}_{(t_i,k_i)} \\
&= n! \sum_{(\mathbf{t}_n,\mathbf{k}_n^{\neg k})\in(\mathbb{X}_T^{\neg t})^{n-1,<}} f_n\big((t_1,k_1),\dots,(t,k),\dots,(t_n,k_n)\big) \prod_{\substack{i=1\\ t_i\neq t}}^{n} \Delta\mathrm{R}_{(t_i,k_i)} \\
&= n! \sum_{(\mathbf{t}_{n-1},\mathbf{k}_{n-1})\in(\mathbb{X}_T^{\neg t})^{n,<}} f_n\big((\mathbf{t}_{n-1},\mathbf{k}_{n-1}),(t,k)\big) \prod_{\substack{i=1\\ t_i\neq t}}^{n} \Delta\mathrm{R}_{(t_i,k_i)} \\
&= n\mathrm{J}_{n-1}\big(f_n(\star,(t,k))\mathbf{1}_{\Delta_n^<}\big) = \mathrm{D}_{(t,k)}\mathrm{J}_n(f_n),
\end{aligned}$$

where $\mathbb{X}_T^{\neg t} = \mathbb{X}_T \backslash \cup\{(t,k),\, k\in\mathrm{E}\}$. Thus, for any $\mathrm{F}\in\mathcal{S}$, $\mathrm{D}_{(t,k)}\mathrm{F} = \big[\mathrm{F}(\pi_t(\eta)+\delta_{(t,k)}) - \mathrm{F}(\pi_t(\eta))\big]$. The result is then extended to $\mathbf{D}$ by a density argument relying on the closability of D (see Corollary 3.9). □

*Proof of Lemma 3.13.* It suffices to state the result for $\mathrm{F} = \xi(h)$, with $h \in \mathrm{L}^2(\mathbb{X})$. By (2.20),

$$\xi(h) = \mathbf{E}\left[\xi(h)\right] + \sum_{m\in\mathbf{N}} \sum_{\substack{\mathrm{J}\subset\mathbf{N}\\ |\mathrm{J}|=m}} \prod_{i\in\mathrm{J}} h(t_i,k_i)\Delta\mathrm{R}_{(t_i,k_i)}.$$

Follows from the definition of $\mathrm{D}^{(n)}$ (3.21) that for any $(\mathbf{t}_n,\mathbf{k}_n)\in\mathbb{X}^n$ and any set $\big\{(t_i,k_i),\, i\in \mathrm{J}, |\mathrm{J}| = m\big\}$ with $m>n$, there exists $i_0\in\mathrm{J}$ such that $(t_{i_0},k_{i_0}) \notin \{(t_i,k_i),\, i\in\{1:n\}\}$. Then, by independence of the $\Delta\mathrm{R}_{(t,k)}$,

$$\mathbf{E}\left[\mathrm{D}^{(n)}_{(\mathbf{t}_n,\mathbf{k}_n)}\Big(\prod_{i\in\mathrm{J}} h(t_i,k_i)\Delta\mathrm{R}_{(t_i,k_i)}\Big)\right] = \mathbf{E}\left[\Delta\mathrm{R}_{(t_{i_0},k_{i_0})}\right]\mathbf{E}\Big[\mathrm{D}^{(n)}_{(\mathbf{t}_n,\mathbf{k}_n)}\Big(\prod_{i\in\mathrm{J}\backslash\{i_0\}} h(t_i,k_i)\Delta\mathrm{R}_{(t_i,k_i)}\Big)\Big].$$

For any $(t,k)\in\mathbb{X}$ let $\mathfrak{r}_{(t,k)}$ be the representative of $\Delta\mathrm{R}_{(t,k)}$. With a similar argument we can prove the same result for any set $\big\{(t_i,k_i),\, i\in\mathrm{J}, |\mathrm{J}|=n\big\}$ different of $(\mathbf{t}_n,\mathbf{k}_n)$ so that

$$\begin{aligned}
\mathbf{E}\Big[\mathrm{D}^{(n)}_{(\mathbf{t}_n,\mathbf{k}_n)}\Big(\prod_{i\in\mathrm{J}} h(t_i,k_i)\Delta\mathrm{R}_{(t_i,k_i)}\Big)\Big] &= \mathbf{E}\Big[\sum_{\mathrm{J}\subset\mathbf{N}} \mathbf{1}_{\{\mathrm{J}=\{1:n\}\}}\Big(\prod_{i\in\mathrm{J}} h(t_i,k_i)\Delta\mathrm{R}_{(t_i,k_i)}\big(\pi_{(\mathbf{t}_n,\mathbf{k}_n)}(\eta)+\delta_{(\mathbf{t}_n,\mathbf{k}_n)}\big) \\
&\qquad - \prod_{i\in\mathrm{J}} h(t_i,k_i)\Delta\mathrm{R}_{(t_i,k_i)}\big(\pi_{(\mathbf{t}_n,\mathbf{k}_n)}(\eta)\big)\Big)\Big] \\
&= \sum_{\mathrm{J}\subset\mathbf{N}} \mathbf{1}_{\{\mathrm{J}=\{1:n\}\}} \prod_{i\in\mathrm{J}} h(t_i,k_i) = n!\prod_{i=1}^{n} h(t_i,k_i) = \mathbf{E}\left[\mathrm{D}^{(n)}_{(\mathbf{t}_n,\mathbf{k}_n)}\mathrm{F}\right].
\end{aligned}$$

On the other hand, by using the alternative characterization of $\mathrm{F}=\xi(h)$, and the orthogonality of the centred variables $\Delta\mathrm{R}$,

$$\mathbf{E}\Big[\mathrm{F}\prod_{i=1}^{n}\frac{1}{\kappa_i}\Delta\mathrm{R}_{(t_i,k_i)}\Big] = \mathbf{E}\Big[\mathbf{E}\left[\xi(h)\right]\prod_{i=1}^{n}\frac{1}{\kappa_i}\Delta\mathrm{R}_{(t_i,k_i)}\Big] + \mathbf{E}\Big[\prod_{s\in\mathbf{N}}\Big(1+\sum_{k\in\mathrm{E}} h(s,k)\Delta\mathrm{R}_{(s,k)}\Big)\prod_{i=1}^{n}\frac{1}{\kappa_i}\Delta\mathrm{R}_{(t_i,k_i)}\Big]$$

$$= \mathbf{E}\left[\prod_{i=1}^{n}\Big(1+\sum_{k\in \mathrm{E}} h(t_i,k)\Delta \mathrm{R}_{(t_i,k)}\Big)\frac{1}{\kappa_i}\Delta \mathrm{R}_{(t_i,k_i)}\right] = \prod_{i=1}^{n} h(t_i,k_i).$$

The result is extended to $\mathrm{L}^2(\mathbf{P})$ by density of the Doléans exponential family. □

*Proof of Lemma 3.14.* It suffices to state the equality for $\mathrm{F}=\xi(f)$, $\mathrm{G}=\xi(g)$, where $f,g\in \mathrm{L}^2(\mathbb{X})$. On the one hand, there exists $T\in\mathbf{N}$ such that

$$\begin{aligned}\mathbf{E}\,[\mathrm{FG}] - \mathbf{E}\,[\mathrm{F}]\,\mathbf{E}\,[\mathrm{G}] &= \prod_{t\in\{1:T\}}\prod_{s\in\{1:T\}}\mathbf{E}\left[\Big(1+\sum_{k\in\mathrm{E}} f(t,k)\Delta\mathrm{R}_{(t,k)}\Big)\Big(1+\sum_{\ell\in\mathrm{E}} g(s,\ell)\Delta\mathrm{R}_{(s,\ell)}\Big)\right]\\ &= \prod_{t\in\{1:T\}}\Big(1+\sum_{k\in\mathrm{E}}\kappa_k f(t,k)g(t,k)\Big)\\ &= \sum_{n\in\{1:T\}}\sum_{\substack{\mathrm{J}\subset\mathbf{N}\\ |\mathrm{J}|=n}}\prod_{j\in\mathrm{J}}\Big(\sum_{k\in\mathrm{E}}\kappa_k f(t_j,k)g(t_j,k)\Big) = \sum_{n\in\{1:T\}}\sum_{\substack{\mathrm{J}\subset\{1:T\}\\ |\mathrm{J}|=n}}\prod_{j\in\mathrm{J}}\langle f(t_j,\cdot),g(t_j,\cdot)\rangle_{\mathrm{L}^2(\mathbb{X},\tilde{\nu})^{\otimes n}}.\end{aligned}$$

On the other hand, for any $n\in\{1:T\}$ and $\mathrm{I}_n\subset\{1:T\}$ of cardinality $n$, denoted $\mathrm{I}_n=\{(t_j^{\mathrm{I}_n},k_j^{\mathrm{I}_n}),\ j\in\{1:n\}\}$ ,

$$\mathbf{E}\left[\mathrm{D}^{(n)}_{\mathrm{I}_n}\mathrm{F}\right] = \prod_{j\in\mathrm{I}_n} f(t_j^{\mathrm{I}_n},k_j^{\mathrm{I}_n})\mathbf{E}\left[\prod_{j\in\{1:T\}\setminus\{1:n\}}\Big(1+\sum_{k\in\mathrm{E}} f(t,k)\Delta\mathrm{R}_{(t,k)}\Big)\right] = \prod_{j\in\mathrm{I}_n} f(t_j,k_j).$$

Then, by denoting by $\mathrm{I}_n^<$ the ordered sets $\mathrm{I}_n$ with respect to the jump times $t_j's$,

$$\sum_{n\in\{1:T\}}\frac{1}{n!}\langle\mathbf{E}[\mathrm{D}^{(n)}\mathrm{F}]\,,\,\mathbf{E}[\mathrm{D}^{(n)}\mathrm{G}]\rangle_{\mathrm{L}^2(\mathbb{X},\tilde{\nu})^{\otimes n}} = \sum_{n\in\{1:T\}}\sum_{\substack{\mathrm{I}_n^<\subset\{1:T\}\\ |\mathrm{I}_n^<|=n}}\prod_{j\in\mathrm{I}_n^<}\langle f(t_j,\cdot),g(t_j,\cdot)\rangle_{\mathrm{L}^2(\mathbb{X},\tilde{\nu})^{\otimes n}}.$$

The result is extended to $\mathrm{L}^2(\mathbf{P})$ by density of the class of Doléans exponentials. □

*Proof of Lemma 3.15.* The proof follows closely that of Last and Penrose ([37], Theorem 1.3). Define for any $\mathrm{F}\in\mathrm{L}^2(\mathbf{P})$ and $n\in\mathbf{Z}_+$ the application $\theta_n^{\mathrm{F}}$ by

$$\theta_n^{\mathrm{F}}(\mathbf{s}_n,\mathbf{l}_n) = \mathbf{E}\left[\mathrm{D}^{(n)}_{(\mathbf{s}_n,\mathbf{l}_n)}\mathrm{F}\right]\ ;\ \forall(\mathbf{s}_n,\mathbf{l}_n)\in\mathbb{X}^n.$$

Let $\mathrm{F}\in\mathrm{L}^2(\mathbf{P})$. The idea is to state the identity for any random variable of the form $\mathrm{G}=\xi(g)$ with $g\in\mathrm{L}^2(\mathbb{X})$, well chosen to approximate F. Indeed follows from the isometry property (2.9) that for $(\mathbf{s}_m,\mathbf{l}_m)\in\mathbb{X}^m$,

$$\mathbf{E}\left[\mathrm{G}\prod_{i=1}^{m}\frac{\Delta\mathrm{R}_{(s_j,\ell_j)}}{\kappa_i}\right] = \mathbf{E}\left[\Big(\mathbf{E}\,[\mathrm{G}] + \sum_{n\in\mathbf{N}}\sum_{(\mathbf{t}_n,\mathbf{k}_n)\in\mathbb{X}^n} g_n(\mathbf{t}_n,\mathbf{k}_n)\prod_{i=1}^{n}\Delta\mathrm{R}_{(t_i,k_i)}\Big)\prod_{j=1}^{m}\frac{\Delta\mathrm{R}_{(s_j,\ell_j)}}{\kappa_j}\right]$$

which is equal to $g_m(\mathbf{s}_m,\mathbf{l}_m)$, whereas by Lemma 3.13, the right member is also equal to $(m!)^{-1}\theta_m^{\mathrm{G}}((\mathbf{s}_m,\mathbf{l}_m))$. Now, from Lemma 3.14 together with the isometry identity (2.6), follows

$$\sum_{n=0}^{\infty}\mathbf{E}\left[\frac{1}{n!}\mathrm{J}_n(f_n)\right]^2 = \sum_{n=0}^{\infty}\frac{1}{n!}\|f_n\|^2_{\mathrm{L}^2(\mathbb{X})^{\circ n}} = \mathbf{E}\left[\mathrm{F}^2\right] < \infty.$$

Hence the infinite series of orthogonal terms $\mathrm{S} := \sum_{n\in\mathbf{Z}_+} \frac{1}{n!}\mathrm{J}_n(\theta_n^{\mathrm{F}})$ converges in $\mathrm{L}^2(\mathbf{P})$. Then,

$$\mathbf{E}\left[(\mathrm{S}-\xi(g))^2\right] = \sum_{n\in\mathbf{Z}_+} \frac{1}{n!}\|\theta_n^{\mathrm{F}} - \theta_n^{\mathrm{G}}\|^2_{\mathrm{L}^2(\mathbb{X})^n} = \mathbf{E}\left[(\mathrm{F}-\xi(g))^2\right],$$

so that, since the set of Doléans exponentials is dense in $\mathrm{L}^2(\mathbf{P})$ and S converges in $\mathrm{L}^2(\mathbf{P})$, the equality $\mathrm{F} = \mathrm{S}$ stands $\mathbf{P}$-almost surely. To prove uniqueness, assume there exists $\mathrm{H} \in \mathrm{L}^2(\mathbf{P})$ which decomposition satisfies (3.12) and such that $\mathrm{S}' := \sum_{n\in\mathbf{Z}_+}(n!)^{-1}\mathrm{J}_n(h_n)$ converges in $\mathrm{L}^2(\mathbf{P})$ to F. Taking the expectation entails $h_0 = \mathbf{E}[\mathrm{F}] = \theta_0^{\mathrm{F}}$. For $n \in \mathbf{N}$, follows from Lemma 3.14 that $\mathbf{E}[\mathrm{F}\mathrm{J}_n(g)] = n!\langle\theta_n^{\mathrm{F}}, g\rangle_{\mathrm{L}^2(\mathbb{X},\tilde{\nu})^{\otimes n}}$ and by replacing $\theta_n^{\mathrm{F}}$ by $h_n$ (since the convergence of $\mathrm{S}'$ holds in $\mathrm{L}^2(\mathbf{P})$) that $\mathbf{E}[\mathrm{F}\mathrm{J}_n(g)] = n!\langle h_n, g\rangle_{\mathrm{L}^2(\mathbb{X},\tilde{\nu})^{\otimes n}}$. Then $\|\theta_n^{\mathrm{F}} - h_n\|^2_{\mathrm{L}^2(\mathbb{X},\tilde{\nu})^{\otimes n}} = \langle\theta_n^{\mathrm{F}} - h_n, g\rangle_{\mathrm{L}^2(\mathbb{X},\tilde{\nu})^{\otimes n}}$ is equal to zero by taking $g = \theta_n^{\mathrm{F}} - h_n$. The proof is thus complete. □

*Proof of Proposition 3.16.* It suffices to prove it for $\mathrm{F} = \xi(h) = \mathfrak{f}(\eta)$ where, as $\eta$ is finite, there exists $T \in \mathbf{N}$ such that

$$\mathfrak{f}(\eta) = \prod_{s\in\{1:T\}} \Big(1 + \sum_{k\in\mathrm{E}} g(s,k)(\mathbf{1}_{\{(s,k)\in\eta\}} - \lambda\mathbf{Q}(\{k\}))\Big).$$

where $g \in \mathrm{L}^2(\mathbb{X}_T)$ is such that $\mathrm{J}_1(h) = \mathrm{J}_1(g\,;\,\mathcal{Z})$. On the one hand, by action of the semi-group P on the quasi-chaotic decomposition (2.14),

$$\mathrm{P}_\tau\mathrm{F} = \xi(e^{-\tau}u) = \prod_{s\in\mathbf{N}} \Big(1 + e^{-\tau}\sum_{k\in\mathrm{E}} g(s,k)(\mathbf{1}_{\{(s,k)\in\eta\}} - \lambda\mathbf{Q}(\{k\}))\Big)$$

On the other hand, by definition of $\eta^{\tau,0}$ (3.22) and $\tilde{\eta}$, which law given $\eta$ is provided by (3.24),

$$\begin{aligned}
\mathbf{E}\left[\mathfrak{f}(\eta^{\tau,0}+\varepsilon\tilde{\eta})\middle|\eta\right] &= \prod_{s\in\{1:T\}} \mathbf{E}\left[1 + \sum_{k\in\mathrm{E}} g(s,k)\big(\mathbf{1}_{\{(s,k)\in(\eta^{\tau,0}+\varepsilon_s^\tau\tilde{\eta})\}} - \lambda\mathbf{Q}(\{k\})\big)\middle|\eta\right]\\
&= \prod_{s\in\{1:T\}} \Big(1 + \sum_{k\in\mathrm{E}} g(s,k)((1-e^{-\tau})\lambda\mathbf{Q}(\{k\})) + e^{-\tau}\mathbf{1}_{\{(s,k)\in\eta\}} - \lambda\mathbf{Q}(\{k\})\Big),\\
&= \prod_{s\in\{1:T\}} \Big(1 + e^{-\tau}\sum_{k\in\mathrm{E}} g(s,k)\big(\mathbf{1}_{\{(s,k)\in\eta\}} - \lambda\mathbf{Q}(\{k\})\big)\Big) = \mathrm{P}_\tau\mathrm{F}.
\end{aligned}$$

Since $\eta$ is finite, the result holds in $\mathrm{L}^2(\mathbf{P})$. The proof is complete. □

*Proof of Proposition 3.17.* Let $\mathrm{F} = \xi(h) = \mathfrak{f}(\eta)$ such that $\mathbf{E}[\mathrm{F}] = 1$ and

$$\mathfrak{f}(\eta) = \prod_{s\in\mathbf{N}} \Big(1 + \sum_{k\in\mathrm{E}} g(s,k)\big(\mathbf{1}_{\{(s,k)\in\eta\}} - \lambda\mathbf{Q}(\{k\})\big)\Big),$$

where $g \in \mathrm{L}^2(\mathbb{X})$ is such that $\mathrm{J}_1(h) = \mathrm{J}_1(g\,;\,\mathcal{Z})$. Then, from Mehler's formula (3.23),

$$\mathrm{P}_\tau\mathfrak{f}(\eta) = \xi(e^{-\tau}h) = \prod_{s\in\mathbf{N}} \Big(1 + e^{-\tau}\sum_{k\in\mathrm{E}} g(s,k)\big(\mathbf{1}_{\{(s,k)\in\eta\}} - \lambda\mathbf{Q}(\{k\})\big)\Big).$$

On the one hand, for any $(s,k) \in \mathbb{X}$,

$$\mathrm{D}_{(s,k)}\mathrm{P}_\tau\mathfrak{f}(\eta) = \prod_{r\in\mathbf{N}}\Big(1+e^{-\tau}\sum_{\ell\in\mathrm{E}} g(r,\ell)\big(\mathbf{1}_{\{(r,\ell)\in(\pi_s(\eta)+\delta_{(s,k)})\}}-\lambda\mathbf{Q}(\{k\})\big)\Big)$$
$$-\prod_{r\in\mathbf{N}}\Big(1+e^{-\tau}\sum_{\ell\in\mathrm{E}} g(r,\ell)\big(\mathbf{1}_{\{(r,\ell)\in\pi_s(\eta)\}}-\lambda\mathbf{Q}(\{k\})\big)\Big) = e^{-\tau}g(s,k)\,\mathrm{P}_\tau\mathfrak{f}(\pi_s(\eta)).$$

On the other hand, follows from

$$\mathrm{D}_{(s,k)}\mathfrak{f}(\eta) = g(s,k)\prod_{r\in\mathbf{N}\setminus\{s\}}\Big(1+\sum_{k\in\mathrm{E}} g(r,k)\big(\mathbf{1}_{\{(r,k)\in\pi_s(\eta)\}}-\lambda\mathbf{Q}(\{k\})\big)\Big) = g(s,k)\mathfrak{f}(\pi_s(\eta)),$$

that for any $(s,k)\in\mathbb{X}$,

$$\mathrm{P}_\tau(\mathrm{D}_{(s,k)}\mathfrak{f}(\pi_s(\eta))) = g(s,k)\prod_{r\in\mathbf{N}\setminus\{s\}}\Big(1+e^{-\tau}\sum_{k\in\mathrm{E}} g(r,k)\big(\mathbf{1}_{\{(r,k)\in\pi_s(\eta)\}}-\lambda\mathbf{Q}(\{k\})\big)\Big)$$
$$= g(s,k)\,\mathrm{P}_\tau\mathfrak{f}(\pi_s(\eta)).$$

Hence the result. □

*Proof of Proposition 3.18.* Let $\mathrm{F}\in\mathrm{L}^2(\mathbf{P})$ such that $\mathbf{E}[\mathrm{F}]=0$. For any $m\in\mathbf{N}$,

$$\mathrm{L}^{-1}\left(\sum_{n=1}^{m}\mathrm{J}_n(f_n)\right) = -\sum_{n=1}^{m}\frac{1}{n}\mathrm{J}_n(f_n) = -\int_0^\infty\sum_{n=1}^{m}e^{-n\tau}\mathrm{J}_n(f_n)\,\mathrm{d}\tau. \tag{6.3}$$

Moreover, the random variable $\mathrm{R}_m$ defined by

$$\mathrm{R}_m := \int_0^\infty\Big(\mathrm{P}_\tau\mathrm{F}-\sum_{n=1}^{m}e^{-n\tau}\mathrm{J}_n(f_n)\Big)\mathrm{d}\tau = \int_0^\infty\Big(\sum_{n=m+1}^{\infty}e^{-n\tau}\mathrm{J}_n(f_n)\Big)\mathrm{d}\tau$$

converges to zero in $\mathrm{L}^2(\mathbf{P})$ by noting that $\mathrm{J}_0(f_0)=\mathbf{E}[\mathrm{F}]=0$ and

$$\mathbf{E}\left[\mathrm{R}_m^2\right]\leqslant\int_0^\infty\mathbf{E}\left[\Big(\sum_{n=m+1}^{\infty}e^{-n\tau}\mathrm{J}_n(f_n)\Big)^2\right]\mathrm{d}\tau = \sum_{n=m+1}^{\infty}n!\|f_n\|^2_{\mathrm{L}^2(\mathbb{X},\tilde{\nu})^{\otimes n}}\int_0^\infty e^{-2n\tau}\,\mathrm{d}\tau.$$

Checking that

$$\mathbf{E}\left[\Big(\int_0^\infty\mathrm{P}_\tau\mathrm{F}\,\mathrm{d}\tau\Big)^2\right]\leqslant\mathbf{E}\left[\int_0^\infty|\mathrm{P}_\tau\mathrm{F}|^2\,\mathrm{d}\tau\right] = \sum_{n=1}^{\infty}n!\|f_n\|^2_{\mathrm{L}^2(\mathbb{X},\tilde{\nu})^{\otimes n}}\int_0^\infty e^{-2n\tau}\,\mathrm{d}\tau$$

is finite, the proof of the first point is complete by letting $n$ go to infinity in (6.3). As for (3.29), the commutation (3.26) and contractivity (3.25) properties satisfied by $(\mathrm{P}_\tau)_{\tau\in\mathbf{R}_+}$ ensure that

$$\mathbf{E}\left[\int_0^\infty|\mathrm{D}_{(s,\ell)}\mathrm{P}_\tau\mathrm{F}|\,\mathrm{d}\tau\right] = \mathbf{E}\left[\int_0^\infty e^{-\tau}|\mathrm{P}_\tau\mathrm{D}_{(s,\ell)}\mathrm{F}|\,\mathrm{d}\tau\right]\leqslant\mathbf{E}\left[|\mathrm{D}_{(s,\ell)}\mathrm{F}|\right]$$

is finite for $\nu$-a.e. $(s,\ell)\in\mathbb{X}$. The result follows by applying the operator D to each side of equality (3.28) and using of the commutation property (3.26). □

### 6.3 Proofs of section 4

*Proof of Theorem 4.1.* Let $\mathbf{Q}$ be an equivalent measure to $\mathbf{P}$ on $\mathcal{F}$. Assume that $\mathrm{E} = \{k^i,\, i \in \mathbf{Z}\}$. Then there exist a real number $\beta$ and a collection of real numbers $\{\alpha_k,\, k \in \mathrm{E}\}$ all in $(0,1)$, satisfying $\beta + \sum_{k\in\mathrm{E}} \alpha_k = 1$, such that

$$\mathbf{Q}_1 = \beta\delta_{\mathbf{0}} + \sum_{k\in\mathrm{E}} \alpha_k \delta_{(1,k)} = \frac{\beta}{1-\lambda}(1-\lambda)\delta_{\mathbf{0}} + \sum_{k\in\mathrm{E}} \frac{\alpha_k}{\lambda\mathbf{Q}(\{k\})}\lambda\mathbf{Q}(\{k\})\delta_{(1,k)} = \mathbf{E}\left[\mathrm{L}_1\mathbf{1}_{\cdot}\right],$$

where $\mathrm{L}_1$ is the random variable defined by

$$\mathrm{L}_1 = \frac{1-\tilde\lambda}{1-\lambda}\mathbf{1}_{\{\eta(\mathbb{X})=0\}} + \sum_{k\in\mathrm{E}} \frac{\tilde\lambda\widetilde{\mathbf{Q}}(\{k\})}{\lambda\mathbf{Q}(\{k\})}\mathbf{1}_{\{(1,k)\in\eta\}} = \frac{1-\tilde\lambda}{1-\lambda} + \sum_{k\in\mathrm{E}}\Big(\frac{\tilde\lambda\widetilde{\mathbf{Q}}(\{k\})}{\lambda\mathbf{Q}(\{k\})} - \frac{1-\tilde\lambda}{1-\lambda}\Big)\mathbf{1}_{\{(1,k)\in\eta\}},$$

with $\tilde\lambda = 1-\beta$ and $\widetilde{\mathbf{Q}}(\{k\}) = \alpha_k/\tilde\lambda$, to get the $\mathbf{Q}(\{k\})$ summoned to 1. Let now L be the $(\mathcal{F}_t)_{t\in\{1:T\}}$-martingale such that

$$\frac{\mathrm{d}\widetilde{\mathbf{P}}}{\mathrm{d}\mathbf{P}}\Big|_{\mathcal{F}_t} = \mathrm{L}_t \;\; ; \; t \in \{1:T\}.$$

Since the increments of the jump process $(\mathrm{N}_t)_{t\in\{1:T\}}$ are independent and identically distributed, we can show by induction that L is defined for any $t \in \{1:T\}$ by

$$\mathrm{L}_t = \prod_{s=1}^{t} \mathrm{L}_s = \prod_{s=1}^{t}\left(\frac{1-\tilde\lambda}{1-\lambda} + \sum_{i\in\mathbf{Z}}\Big(\frac{\tilde\lambda\widetilde{\mathbf{Q}}(\{k^i\})}{\lambda\mathbf{Q}(\{k^i\})} - \frac{1-\tilde\lambda}{1-\lambda}\Big)\mathbf{1}_{\{(s,k^i)\in\eta\}}\right).$$

Let $u$ be the process defined by (4.1). For any $t \in \{1:T\}$, by using $\mathbf{1}_{\{(s,k^i)\in\eta\}} = \Delta\mathrm{Z}_{(s,k^i)} + \lambda\mathbf{Q}(\{k^i\})$ and noting that

$$\sum_{i\in\mathbf{Z}}\Big(\frac{\tilde\lambda\widetilde{\mathbf{Q}}(\{k^i\})}{\lambda\mathbf{Q}(\{k^i\})} - \frac{1-\tilde\lambda}{1-\lambda}\Big)\lambda\mathbf{Q}(\{k^i\}) = \tilde\lambda - \frac{\lambda(1-\tilde\lambda)}{1-\lambda} = 1 - \frac{1-\tilde\lambda}{1-\lambda},$$

we can check that, by the definition of $h$ given in the theorem, that

$$\xi_t(h) = \prod_{s=1}^{t}\Big(1+\sum_{i\in\mathbf{Z}} g(s,k^i)\Delta\mathrm{Z}_{(s,k^i)}\Big) = \prod_{s=1}^{t}\left(\frac{1-\tilde\lambda}{1-\lambda} + \sum_{i\in\mathbf{Z}}\Big(\frac{\tilde\lambda\widetilde{\mathbf{Q}}(\{k^i\})}{\lambda\mathbf{Q}(\{k^i\})} - \frac{1-\tilde\lambda}{1-\lambda}\Big)\mathbf{1}_{\{(s,k^i)\in\eta\}}\right) = \mathrm{L}_t.$$

Hence the result. □

*Proof of Corollary 4.2.* Let $\varphi$ and $\widetilde{\mathbf{P}}$ as defined in the theorem. Let $\widetilde{\mathbf{E}}$ denote the expectation taken under the probability measure $\widetilde{\mathbf{P}}$. For any $s \in \mathbf{R}^*$, $t \in \mathbf{N}$,

$$\begin{aligned}\widetilde{\mathbf{E}}[s^{\mathrm{Y}_t}] &= \left(\frac{1-\tilde\lambda}{1-\lambda}\right)^t \sum_{n=0}^{t} \mathbf{E}\left[s^{\mathrm{Y}_t}\prod_{r=1}^{n}(1+\varphi(\mathrm{V}_r))\Big|\mathrm{N}_t = n\right]\mathbf{P}(\mathrm{N}_t = n) \\ &= \left(\frac{1-\tilde\lambda}{1-\lambda}\right)^t \sum_{n=0}^{t}\binom{t}{n}\lambda^n(1-\lambda)^{t-n}\mathbf{E}\left[\prod_{r=1}^{n}(1+\varphi(\mathrm{V}_r))s^{\mathrm{V}_r}\Big|\mathrm{N}_t = n\right]\end{aligned}$$

$$
\begin{aligned}
&= (1-\tilde\lambda)^t \sum_{n=0}^{t} \binom{t}{n} \left(\frac{\lambda}{1-\lambda}\right)^n \left(\frac{\tilde\lambda(1-\lambda)}{\lambda(1-\tilde\lambda)}\right)^n \prod_{r=1}^{n} \left(\sum_{k\in \mathrm{E}} \frac{\widetilde{\mathbf{Q}}(\{k\})}{\mathbf{Q}(\{k\})} \cdot \mathbf{Q}(\{k\}) s^k\right)\\
&= \sum_{n=0}^{t} \binom{t}{n} \tilde\lambda^n (1-\tilde\lambda)^{t-n} \Big(\sum_{k\in\mathrm{E}} \widetilde{\mathbf{Q}}(\{k\}) s^k\Big)^n = \Big(1-\tilde\lambda + \tilde\lambda \widetilde{\mathbf{E}}_{\mathbf{Q}}[s^{\mathrm{V}}]\Big)^t,
\end{aligned}
$$

where V is a E-valued random variable and $\widetilde{\mathbf{E}}_{\mathbf{Q}}$ is the expectation taken under the probability measure $\widetilde{\mathbf{Q}}$. Hence the result. □

*Proof of Theorem 4.4.* Let $\mathrm{F} \in \mathcal{S}$; follows from both its chaotic decomposition (2.14) and the definition of the gradient operator (3.13) that for some $T \in \mathbf{N}$,

$$
\begin{aligned}
\mathrm{F} &= \mathbf{E}[\mathrm{F}] + \sum_{n\in\mathbf{N}} \mathrm{J}_n(f_n \mathbf{1}_{\{1:T\}^n})\\
&= \mathbf{E}[\mathrm{F}] + \sum_{n\in\mathbf{N}} n \sum_{(t,k)\in\mathbb{X}_T} \mathrm{J}_{n-1}\big(f_n(\star,(t,k))\mathbf{1}_{\{1:t-1\}^{n-1,<}}\big)\,\Delta\mathrm{R}_{(t,k)}\\
&= \mathbf{E}[\mathrm{F}] + \sum_{(t,k)\in\mathbb{X}_T} \sum_{n\in\mathbf{N}} \mathbf{E}\left[n\mathrm{J}_{n-1}\big(f_n(\star,(t,k))\big|\mathcal{F}_{t-1}\right]\,\Delta\mathrm{R}_{(t,k)}\\
&= \mathbf{E}[\mathrm{F}] + \sum_{(t,k)\in\mathbb{X}_T} \mathbf{E}\left[\mathrm{D}_{(t,k)}\mathrm{F}\,|\,\mathcal{F}_{t-1}\right]\,\Delta\mathrm{R}_{(t,k)},
\end{aligned}
$$

where we have used lemma 2.8 to get the third line. As noticed in Remark 4.5, the operator $\mathrm{F} \in \mathrm{L}^2(\mathbf{P}) \mapsto \big(\mathbf{E}[\mathrm{D}_{(t,k)}\mathrm{F}],\,(t,k)\in\mathbb{X}\big)$ is bounded with norm equal to 1; the result can be thus extended to any random variable $\mathrm{F}\in\mathrm{L}^2(\mathbf{P})$ using a standard Cauchy argument. □

*Proof of Corollary 4.7.* According to (4.3), we have

$$
\begin{aligned}
\operatorname{var}(\mathrm{F}) &= \mathbf{E}\left[\left|\sum_{(t,k)\in\mathbb{X}} \mathbf{E}\left[\mathrm{D}_{(t,k)}\mathrm{F}\,|\,\mathcal{F}_{t-1}\right]\,\Delta\mathrm{R}_{(t,k)}\right|^2\right]\\
&= \mathbf{E}\left[\sum_{(t,k)\in\mathbb{X}} \kappa_k \Big|\mathbf{E}\left[\mathrm{D}_{(t,k)}\mathrm{F}\,|\,\mathcal{F}_{t-1}\right]\Big|^2\right]\\
&\leqslant \mathbf{E}\left[\sum_{(t,k)\in\mathbb{X}} \kappa_k \mathbf{E}\left[|\mathrm{D}_{(t,k)}\mathrm{F}|^2|\mathcal{F}_{t-1}\right]\right] = \mathbf{E}\left[\int_{\mathbb{X}} |\mathrm{D}_{(t,k)}\mathrm{F}|^2\,\mathrm{d}\tilde\nu(t,k)\right],
\end{aligned}
$$

where the inequality follows from Jensen's. □

## 6.4 Proofs of section 5

### 6.4.1 Proofs of Subsection 5.1

*Proof of Theorem 5.1.* Let $\mathrm{F}\in\mathbf{D}$ be a $\mathbf{Z}_+$-valued random variable such that $\mathbf{E}[\mathrm{F}] = \lambda_0$. Let $\mathrm{A}\subset\mathbf{Z}_+$ and $\varphi_\mathrm{A}$ be the solution of the Stein equation (5.3). The uniform boundedness of $\nabla\varphi_\mathrm{A}$ ensures that $\varphi_\mathrm{A}(\mathrm{F})\in\mathbf{D}$, whereas

$$
\big|\varphi_\mathrm{A}(\mathrm{F})\mathrm{D}^+\mathrm{L}^{-1}(\mathrm{F}-\mathbf{E}[\mathrm{F}])\big| \leqslant \|\varphi\|_\infty |\mathrm{D}^+\mathrm{L}^{-1}(\mathrm{F}-\mathbf{E}[\mathrm{F}]))|
$$

implies that $\varphi_A(F)D^+L^{-1}(F-\mathbf{E}[F]) \in L^1(\mathbf{P}\otimes\nu)$. Under this condition together with the hypotheses of the theorem, the combination of $L^1$ and $L^2$ theories performs; in particular, since $F \in \mathbf{D}$, $DF = D^+F$ for all $F \in L^2(\mathbf{P})$, almost surely. Moreover, as $E = \{1\}$, follows from remark 3.20 that $\widetilde{L}F = LF$ $\mathbf{P}$-almost surely and that the integration by parts formula (3.18) holds. Note besides that here $\pi_t(\eta)+\delta_t = \eta+\delta_t$ so that $\widetilde{D}_tF = \mathfrak{f}(\eta+\delta_t)-\mathfrak{f}(\eta)$, $\mathbf{P}$-almost surely. By definition of the operators L and $L^{-1}$ and using (3.10), we get

$$\begin{aligned}\mathbf{E}[F\varphi_A(F)-\lambda_0\varphi_A(F+1)] &= \mathbf{E}[(F-\mathbf{E}[F])\varphi_A(F)-\lambda_0\nabla\varphi_A(F)]\\ &= \mathbf{E}\left[(\widetilde{L}\widetilde{L}^{-1}(F-\mathbf{E}[F]))\varphi_A(F)\right]-\mathbf{E}[\lambda_0\nabla\varphi_A(F)]\\ &= -\mathbf{E}\left[\langle\widetilde{D}(\varphi_A(F)), DL^{-1}(F-\mathbf{E}[F])\rangle_{L^2(\mathbb{X},\nu)}\right]-\mathbf{E}[\lambda_0\nabla\varphi_A(F)]\\ &= -\mathbf{E}\left[\nabla\varphi_A(F)\langle\widetilde{D}F, DL^{-1}(F-\mathbf{E}[F])\rangle_{L^2(\mathbb{X},\nu)}+\mathrm{rem}\right]-\mathbf{E}[\lambda_0\nabla\varphi_A(F)],\end{aligned}$$

where we have used to get the third line that

$$\begin{aligned}\langle\widetilde{D}(\varphi_A(F)), DL^{-1}(F-\mathbf{E}[F])\rangle_{L^2(\mathbb{X},\nu)} = \nabla\varphi_A(F)\int_{\mathbf{N}}\left(\widetilde{D}_tF\right)\left(D_tL^{-1}(F-\mathbf{E}[F])\right)\nu(\mathrm{d}t)\\ +\int_{\mathbf{N}}\mathfrak{R}_t\left(D_tL^{-1}F\right)\nu(\mathrm{d}t),\end{aligned}$$

where $\mathfrak{R}_t$ is a residual random function such that $\mathfrak{R}_t \leqslant \|\nabla^2\varphi_A\|_\infty\left|(\widetilde{D}_tF)(\widetilde{D}_tF-1)\right|/2$. By using inequality (5.2), with $k=\mathfrak{f}(\eta)+\widetilde{D}_tF$ and $a=\mathfrak{f}(\eta)$, we get

$$\begin{aligned}\mathbf{E}[|\mathrm{rem}|] &= \mathbf{E}\left[\left|\int_{\mathbf{N}}\mathfrak{R}_t(D_tL^{-1}(F-\mathbf{E}[F]))\,\nu(\mathrm{d}t)\right|\right]\\ &\leqslant \frac{\|\nabla^2\varphi_A\|_\infty}{2}\mathbf{E}\left[\int_{\mathbf{N}}|(\widetilde{D}_tF)(\widetilde{D}_tF-1)|\left|D_tL^{-1}(F-\mathbf{E}[F])\right|\nu(\mathrm{d}t)\right].\end{aligned}$$

Then,

$$\begin{aligned}\left|\mathbf{E}[F\varphi_A(F)-\lambda_0\varphi_A(F+1)]\right| &\leqslant \|\nabla\varphi_A\|_\infty\mathbf{E}\left[\left|\lambda_0-\langle\widetilde{D}F, -DL^{-1}(F-\mathbf{E}[F])\rangle_{L^2(\mathbb{X},\nu)}\right|\right]\\ &+\frac{\|\nabla^2\varphi_A\|_\infty}{2}\mathbf{E}\left[\int_{\mathbf{N}}\left|(\widetilde{D}_tF)(\widetilde{D}_tF-1)\right|\left|D_tL^{-1}(F-\mathbf{E}[F])\right|\nu(\mathrm{d}t)\right].\end{aligned}$$

The result is then stated by taking the supremum over the set $\{A \subset \mathbf{Z}_+\}$ and using the uniform bounds (over the class ) estimates on $\nabla.\varphi$ and $\nabla^2_.\varphi$. □

*Proof of Proposition 5.2.* Let $F \in \mathbf{D}$ be a $\mathbf{Z}_+$-valued random variable such that $\mathbf{E}[F] = \lambda_0\mathbf{E}[V_1]$. Via Stein's method and the definition of the Stein operator for Poisson compound approximation, we are led to control

$$\mathbf{E}\Big[F\psi_A(F)-\int_{\mathbb{X}}k\psi_A(F+k)\mathrm{d}\nu(t,k)\Big] = \mathbf{E}\Big[(F-\mathbf{E}[F])\psi_A(F)-\int_{\mathbb{X}}k(\psi_A(F+k)-\psi_A(F))\mathrm{d}\nu(t,k)\Big],$$

where, by definition of the operators $\widetilde{L}$ and $\widetilde{L}^{-1}$, we have

$$\mathbf{E}[(F-\mathbf{E}[F])\psi_A(F)] = \mathbf{E}\big[(\widetilde{L}\widetilde{L}^{-1})(F-\mathbf{E}[F])\psi_A(F)\big] = \mathbf{E}\Big[\langle D\widetilde{L}^{-1}(F-\mathbf{E}[F]), \widetilde{D}\psi_A(F)\rangle_{L^2(\mathbb{X},\nu)}\Big].$$

On the other hand,

$$\begin{aligned}
\langle \mathrm{D}^+\widetilde{\mathrm{L}}^{-1}(\mathrm{F}-\mathbf{E}[\mathrm{F}]), \widetilde{\mathrm{D}}\psi_{\mathrm{A}}(\mathrm{F})\rangle_{\mathrm{L}^2(\mathbb{X},\nu)} &- \int_{\mathbb{X}} k(\psi_{\mathrm{A}}(\mathrm{F}+k)-\psi_{\mathrm{A}}(\mathrm{F}))\,\mathrm{d}\nu(t,k)\\
&= \int_{\mathbb{X}} \mathrm{D}^+\widetilde{\mathrm{L}}^{-1}\big(\mathfrak{f}(\eta)-\mathbf{E}[\mathfrak{f}(\eta)]\big)\big[\psi_{\mathrm{A}}\big(\mathfrak{f}(\pi_t(\eta)+\delta_{(t,k)})\big)-\psi_{\mathrm{A}}(\mathfrak{f}(\eta))\big]\,\mathrm{d}\nu(t,k)\\
&\qquad - \int_{\mathbb{X}} k\big[\psi_{\mathrm{A}}(\mathfrak{f}(\eta)+k)-\psi_{\mathrm{A}}(\mathfrak{f}(\eta))\big]\,\mathrm{d}\nu(t,k)\\
&= \int_{\mathbb{X}} \big[\mathrm{D}^+\widetilde{\mathrm{L}}^{-1}\big(\mathfrak{f}(\eta)-\mathbf{E}[\mathfrak{f}(\eta)]\big)\psi_{\mathrm{A}}\big(\mathfrak{f}(\pi_t(\eta)+\delta_{(t,k)})\big)-k\psi_{\mathrm{A}}(\mathfrak{f}(\eta)+k)\big]\,\mathrm{d}\nu(t,k)\\
&\qquad - \int_{\mathbb{X}} \psi_{\mathrm{A}}(\mathfrak{f}(\eta))\big[\mathrm{D}^+\widetilde{\mathrm{L}}^{-1}\big(\mathfrak{f}(\eta)-\mathbf{E}[\mathfrak{f}(\eta)]\big)-k\big]\,\mathrm{d}\nu(t,k).
\end{aligned}$$

The conclusion follows by taking the expectation and then the supremum over the class $\{\psi_{\mathrm{A}},\ \mathrm{A}\subset\mathbf{Z}_+\}$. ☐

*Proof of Theorem 5.4.* Since U is a simple binomial functional (since the space mark boils down to a singleton), $\mathrm{L}^1$ and $\mathrm{L}^2$ theories combine perfectly, $\mathrm{DU}=\mathrm{D}^+\mathrm{U}$, $\widetilde{\mathrm{D}}_t\mathrm{U}=\mathbf{1}_{\{t\notin\eta\}}\mathrm{D}^+\mathrm{U}$, and we get $\widetilde{\mathrm{L}}\mathrm{U}=\mathrm{LU}$ $\mathbf{P}$-almost surely. Note that $\lambda_0^2=p^{2m}+2(n-1)qp^{2m}+(n-1)^2q^2p^{2m}$. Then,

$$\begin{aligned}
\mathbf{E}\big[\langle\widetilde{\mathrm{D}}\mathrm{U},-\mathrm{DL}^{-1}(\mathrm{U}-\mathbf{E}[\mathrm{U}])\rangle_{\mathrm{L}^2(\mathbb{X},\nu)}\big] &= \mathbf{E}\big[(\mathrm{LL}^{-1}(\mathrm{U}-\mathbf{E}[\mathrm{U}]))\mathrm{U}\big]=\mathrm{var}[\mathrm{U}]\\
&= \lambda_0+2\sum_{i=m+1}^{n-1}\mathbf{E}[\mathrm{U}_0\mathrm{U}_i]+2\sum_{1\leqslant i<j}\mathbf{E}[\mathrm{U}_i\mathrm{U}_j]-\lambda_0^2\\
&= \lambda_0+2(n-m-1)qp^{2m}+(n-1)(n-2m-2)q^2p^{2m}-\lambda_0^2\\
&= \lambda_0-2mqp^{2m}-(2m-1)q^2p^{2m}-p^{2m}. \qquad (6.4)
\end{aligned}$$

Then,

$$\big|\lambda_0-\mathbf{E}\big[\langle\widetilde{\mathrm{D}}\mathrm{U},-\mathrm{DL}^{-1}(\mathrm{U}-\mathbf{E}[\mathrm{U}])\rangle_{\mathrm{L}^2(\mathbb{X},\nu)}\big]\big| = p^{2m}[2(m-1)q^2+2mq+1].$$

On the other hand, using Corollary 3.18

$$\begin{aligned}
\mathbf{E}\Big[\int_{\mathbf{N}}\big|(\widetilde{\mathrm{D}}_s\mathrm{U})(\widetilde{\mathrm{D}}_s\mathrm{U}-1)\big|\big|\mathrm{D}_s\mathrm{L}^{-1}(\mathrm{U}-\mathbf{E}[\mathrm{U}])\big|\nu(\mathrm{d}s)\Big]&\\
\leqslant \mathbf{E}\left[\int_{\{1:n\}}\Big(\mathbf{1}_{\{t\notin\eta\}}\big|(\mathrm{D}_s\mathrm{U})^2-(\mathrm{D}_s\mathrm{U})\big|\Big|\int_0^\infty \mathrm{D}_s\mathrm{P}_\tau\mathrm{U}\,\mathrm{d}\tau\Big|\Big)\nu(\mathrm{d}s)\right]&\\
= \mathbf{E}\left[\int_{\{1:n\}}\Big(\big|(\mathrm{D}_s\mathrm{U})^2-(\mathrm{D}_s\mathrm{U})\big|\Big|e^{-s}\mathrm{D}_s\mathrm{U}\,\mathrm{d}\tau\Big|\Big)\nu(\mathrm{d}s)\right]&,
\end{aligned}$$

where we have used that $\mathrm{D}_t\mathrm{U}$ is $\mathcal{G}_t$-measurable. Moreover,

$$(\mathrm{D}_t\mathrm{U})^3=\Bigg(\prod_{i=1,i\neq t}^{m}\Delta\mathrm{N}_i+\sum_{i=t-m}^{t-1}(1-\Delta\mathrm{N}_i)\prod_{\ell=1,i+\ell\neq t}^{m}\Delta\mathrm{N}_{i+\ell}-\prod_{\ell=1}^{m}\Delta\mathrm{N}_{t+\ell}\Bigg)^3=:(\mathrm{A}+\mathrm{B}-\mathrm{C})^3,$$

Note first that for $t>m$, $\mathrm{D}_t\big(\prod_{i=1}^m\Delta\mathrm{N}_i\big)=0$ so that $\mathrm{D}_t\mathrm{U}=\mathrm{B}-\mathrm{C}$. Assume there exists

$i_0 \in \{t-m, t-1\}$ such that $\mathrm{B}_{i_0} = 1$ where $\mathrm{B}_i := (1-\Delta\mathrm{N}_i)\prod_{\ell=1, i+\ell\neq t}^{m} \Delta\mathrm{N}_{i+\ell}$. Then, $\Delta\mathrm{N}_{i_0} = 0$ implies $\mathrm{B}_i = 0$ for any $i \in \{\max(0, i_0 - m), i_0 - 1\}$ whereas $\Delta\mathrm{N}_{i_0+\ell} = 1$ leads to $\mathrm{B}_i = 0$ for any $i \in \{i_0 + 1 : \min(i_0 + m, t-1)\}$; this entails $\mathrm{B} = \sum_{i=t-m}^{t-1} \mathrm{V}_i = 0 + \mathrm{B}_{i_0} = 1$. Thus $\mathrm{B} \in \{0, 1\}$ and $\mathbf{P}(\{\mathrm{B} \leqslant 1\}) = 1$. Besides,

$$\begin{aligned}(\mathrm{D}_t\mathrm{U})^3 &= \mathrm{B}^3 - 3\mathrm{B}^2\mathrm{C} + 3\mathrm{B}\mathrm{C}^2 - \mathrm{C}^3\\ &= (\mathrm{B} - 2\mathrm{BC} + \mathrm{C}) + 2\mathrm{BC} - 2\mathrm{C} = (\mathrm{D}_t\mathrm{U})^2 + 2\mathrm{C}(\mathrm{B} - 1).\end{aligned}$$

Since $\mathrm{B}, \mathrm{C} \in \{0, 1\}$, this proves that $(\mathrm{D}_t\mathrm{U})^3 \leqslant (\mathrm{D}_t\mathrm{U})^2$ $\mathbf{P}$-p.s. On the event $\{\mathrm{C} = 1\}$, $\prod_{\ell=1}^{m} \Delta\mathrm{N}_{t+\ell} = 1$ so that $\mathrm{B}_{t-1}$ is equal to 0 provided $\Delta\mathrm{N}_{t-1} = 1$. This entails $\prod_{\ell=1,\ell\neq 2}^{m} \Delta\mathrm{N}_{t-2+\ell}$ is equal to 1 and then $\mathrm{B}_{t-2}$ is equal to 0 if and only if $\Delta\mathrm{N}_{t-2} = 1$. By induction, we can prove that, on $\{\mathrm{C} = 1\}$, $\mathrm{B} = 0$ if and only if $\Delta\mathrm{N}_i = 1$ for any $i \in \{t-m, \ldots, t-1\}$ (with probability $p^m$). Then

$$\mathbf{E}\left[\mathrm{C}(1 - \mathrm{B})\right] = \mathbf{E}\left[\mathbf{1}_{\{\mathrm{B}=0\}} \middle| \{\mathrm{C} = 1\}\right] \mathbf{P}(\{\mathrm{C} = 1\}) = p^{2m},$$

so that we get for $t > m$,

$$\mathbf{E}\left[(\mathrm{D}_t\mathrm{U})^2 - (\mathrm{D}_t\mathrm{U})^3\right] = p^{2m},$$

On the other hand for $t \leqslant m$, $\mathrm{D}_t\mathrm{U} = (\mathrm{A}+\mathrm{B}+\mathrm{C})$ and $\mathrm{AB} = 0$. In that case, since $\mathrm{A}, \mathrm{C} \in \{0, 1\}$,

$$\mathbf{E}\left[((\mathrm{D}_t\mathrm{U})^3 - (\mathrm{D}_t\mathrm{U})^2)\mathbf{1}_{\{\mathrm{A}=1\}}\right] = \mathbf{E}\left[(1 - \mathrm{C})^2\mathrm{C}\mathbf{1}_{\{\mathrm{A}=1\}}\right] = 0.$$

On $\{\mathrm{A} = 0\}$, $\mathrm{D}_t\mathrm{U} = \mathrm{B} - \mathrm{C}$ and we can reason as above to prove that for $t \leqslant m$, $\mathbf{E}\left[((\mathrm{D}_t\mathrm{U})^3 - (\mathrm{D}_t\mathrm{U})^2)\right] = (1 - p^{m-1})\mathbf{E}\left[((\mathrm{D}_t\mathrm{U})^3 - (\mathrm{D}_t\mathrm{U})^2)\mathbf{1}_{\{\mathrm{A}=0\}}\right] = (1 - p^{m-1})p^{2m} \leqslant p^{2m}$. Finally,

$$\begin{aligned}&\int_{\{1:n-m-1\}} \mathbf{E}\Big[\big|(\widetilde{\mathrm{D}}_t\mathrm{U})(\widetilde{\mathrm{D}}_t\mathrm{U} - 1)\big|\big|\mathrm{D}_t\mathrm{L}^{-1}(\mathrm{U} - \mathbf{E}\left[\mathrm{U}\right])\big|\Big]\, \nu(\mathrm{d}t)\\ &= \int_{\{1:n-m-1\}} \mathbf{E}\Big[\mathbf{1}_{\{t\notin\eta\}}\big|(\mathrm{D}_t\mathrm{U})(\mathrm{D}_t\mathrm{U}-1)\big|\big|\mathrm{D}_t\mathrm{L}^{-1}(\mathrm{U}-\mathbf{E}\left[\mathrm{U}\right])\big|\Big]\, \nu(\mathrm{d}t) \leqslant (n-m-1)(1-p)p^{2m+1}.\end{aligned}$$

This together with (6.4) provides the bound in Theorem 5.2. □

*Proof of Theorem 5.7.* Consider the random variable $\mathrm{H} = \sum_{j\in\mathcal{J}} \mathrm{V}_j\Delta\mathrm{N}_j$, as defined in the theorem. Since H belongs to $\mathcal{H}_1$, then $\widetilde{\mathrm{L}}\mathrm{H} = \mathrm{H}$ so that $\mathrm{D}\widetilde{\mathrm{L}}^{-1}(\mathrm{H}-\mathbf{E}\left[\mathrm{H}\right]) = \mathrm{D}^+\widetilde{\mathrm{L}}^{-1}(\mathrm{H}-\mathbf{E}\left[\mathrm{H}\right]) = \mathrm{D}^+\mathrm{H}$. Moreover for any $(t, k) \in \mathcal{J} \times \mathbf{N}$, $\mathrm{D}^+_{(t,k)}\mathrm{H} = k$ and $\widetilde{\mathrm{D}}_{(t,k)}\mathrm{H} = (k - \ell)\mathbf{1}_{\{(t,\ell)\in\eta\}}$ $\mathbf{P}$-almost surely. Then the second term in (5.7) vanishes and it remains to control

$$\int_{\mathbb{X}} k\mathbf{E}\left[\psi_{\mathrm{A}}(\mathfrak{h}(\pi_t(\eta) + \delta_{(t,k)})) - \psi_{\mathrm{A}}(\mathfrak{h}(\eta) + k)\right] \mathrm{d}\nu(t, k).$$

On the other hand, by denoting $\mathrm{H}^{\neg t} = \sum_{j\in\mathcal{J}\setminus\{t\}} \mathrm{V}_j\Delta\mathrm{N}_j$,

$$\begin{aligned}&\mathbf{E}\big[\psi_{\mathrm{A}}(\mathfrak{h}(\pi_t(\eta)+\delta_{(t,k)})) - \psi_{\mathrm{A}}(\mathfrak{h}(\eta) + k)\big]\\ &\qquad= \sum_{\ell\in\mathbf{N}} \mathbf{E}\left[\Big(\psi_{\mathrm{A}}\big(\mathfrak{h}(\pi_t(\eta) + \delta_{(t,k)})\big) - \psi_{\mathrm{A}}\big(\mathfrak{h}(\eta) + k\big)\Big)\mathbf{1}_{\{(t,\ell)\in\eta\}}\right]\\ &\qquad= \sum_{\ell\in\mathbf{N}} \mathbf{E}\left[\Big(\psi_{\mathrm{A}}\big(\mathfrak{h}(\pi_t(\eta) + \delta_{(t,k)})\big) - \psi_{\mathrm{A}}\big(\mathfrak{h}(\pi_t(\eta) + \delta_{(t,\ell)}) + k\big)\Big)\mathbf{1}_{\{(t,\ell)\in\eta\}}\right]\end{aligned}$$

$$\leqslant \|\nabla\psi_{\mathrm{A}}\|_\infty \sum_{\ell\in\mathbf{N}} \mathbf{E}\left[\Big(\mathrm{H}^{\neg t}+k-\big(\mathrm{H}^{\neg t}+k+\ell\big)\Big)\mathbf{1}_{\{(t,\ell)\in\eta\}}\right]$$
$$= \|\nabla\psi_A\|_\infty \sum_{\ell\in\mathbf{N}} \ell\mu(\mathrm{W})(1-\alpha)^2\alpha^{\ell-1} = \|\nabla\psi_A\|_\infty\mu(\mathrm{W}).$$

Then,

$$\Big|\int_{\mathbb{X}} k\mathbf{E}\left[\psi_{\mathrm{A}}[\mathfrak{h}(\pi_t(\eta)+\delta_{(t,k)})]-\psi_{\mathrm{A}}[\mathfrak{h}(\eta)+k]\right]\,\mathrm{d}\nu(t,k)\Big| \leqslant (n-h+1)\mathfrak{d}_{\mathcal{PC}}\mu(\mathrm{W})^2.$$

This provides a bound for $\mathrm{dist}_{\mathrm{TV}}(\mathbf{P}_{\mathrm{H}},\mathcal{PC}(\lambda_0,\mathbf{V}))$ and the conclusion follows by using the triangular inequality together with the previous bounds. □

### 6.4.2 Proofs of subsection 5.2

*Proof of Lemma 5.14.* From Schweizer ([70], proof of Lemma 2.7), $\xi_t^{\mathrm{F}}$ can be simply written

$$\xi_t^{\mathrm{F}} = \frac{\mathbf{E}\left[\Delta\widehat{\mathbf{E}}\left[\mathrm{F}|\mathcal{F}_t\right]\Delta\mathrm{S}_t\big|\mathcal{F}_{t-1}\right]}{\mathbf{E}\left[(\Delta\mathrm{S}_t)^2\,|\,\mathcal{F}_{t-1}\right]}\ ;\ t\in\mathbf{T}.$$

The application of the Clark decomposition to $\widehat{\mathbf{E}}\left[\mathrm{F}|\mathcal{F}_t\right]-\widehat{\mathbf{E}}\left[\mathrm{F}|\mathcal{F}_{t-1}\right]$ yields

$$\xi_t^{\mathrm{F}} = \frac{1}{\mathbf{E}\left[(\Delta\mathrm{S}_t)^2\,|\,\mathcal{F}_{t-1}\right]}\left(\sum_{k\in\mathrm{E}}\mathbf{E}\Big[\widehat{\mathbf{E}}\big[\mathrm{D}_{(t,k)}\mathrm{F}|\mathcal{F}_{t-1}\big]\Delta\mathrm{R}_{(t,k)}\mathrm{S}_{t-1}\big((b-a\rho)\Delta\mathrm{R}_{(t,1)}+a\Delta\mathrm{R}_{(t,-1)}+\lambda r\big)\,\Big|\,\mathcal{F}_{t-1}\Big]\right)$$
$$= \frac{1}{\mathrm{S}_{t-1}v_t}\sum_{k\in\mathrm{E}} u_t\widehat{\mathbf{E}}\big[\mathrm{D}_{(t,k)}\mathrm{F}|\mathcal{F}_{t-1}\big],$$

where we have used that $r=bp+aq$ and that $\mathbf{E}\left[\Delta\mathrm{R}_{(t,\ell)}\Delta\mathrm{R}_{(t,k)}|\mathcal{F}_{t-1}\right]=0$ for $\ell\neq k$ due to the orthogonality of the family $\mathcal{R}$. The sequences $u=(u_t)_{t\in\mathbf{T}}$ and $v=(v_t)_{t\in\mathbf{T}}$ are defined by $v_t=\lambda(b^2p+a^2q)$ and

$$u_t = \lambda\big[(b-a\rho)\mathbf{E}\left[(\Delta\mathrm{R}_{(t,1)})^2\right]+a\mathbf{E}\left[(\Delta\mathrm{R}_{(t,-1)})^2\right]\big]$$
$$= \lambda\big[(b-a\rho)p(1-\lambda p)+a\big(q(1-\lambda q)-\rho\lambda pq\big)\big].$$

Hence the result. □

**Acknowledgements.** This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N°811017. I am also grateful to Giovanni Peccati for motivating discussions.

## References

[1] R. Arratia, L. Goldstein, and L. Gordon. Two moments suffice for Poisson approximations: the Chen-Stein method. *The Annals of Probability*, 17(1):9–25, 1989.

[2] R. Arratia, L. Goldstein, and L. Gordon. Poisson approximation and the Chen-Stein method. *Statistical Science*, pages 403–424, 1990.

[3] E. Azmoodeh, S. Campese, and G. Poly. Fourth moment theorems for Markov diffusion generators. *Journal of Functional Analysis*, 266(4):2341–2359, Feb 2014.

[4] E. Azmoodeh, D. Malicet, G. Mijoule, and G. Poly. Generalization of the Nualart-Peccati criterion. *Annals of Probability*, 44(2):924–954, 2016.

[5] A. D. Barbour and L. H. Y. Chen. *An introduction to Stein's method*, volume 4 of *Lecture Notes Series*. National University of Singapore, 2005.

[6] A. D. Barbour, L. H. Y. Chen, and W.-L. Loh. Compound Poisson approximation for nonnegative random variables via Stein's method. *The Annals of Probability*, pages 1843–1866, 1992.

[7] A. D. Barbour and O. Chryssaphinou. Compound Poisson approximation: a user's guide. *Annals of Applied Probability*, pages 964–1002, 2001.

[8] A. D. Barbour, L. Holst, and S. Janson. *Poisson approximation*, volume 2. The Clarendon Press Oxford University Press, 1992.

[9] A. D. Barbour and S. Utev. Solving the Stein equation in compound Poisson approximation. *Advances in Applied Probability*, pages 449–475, 1998.

[10] K. Bichteler. Malliavin calculus for processes with jumps. *Stochastics Monographs*, 1987.

[11] S. Bourguin and G. Peccati. The Malliavin-Stein method on the Poisson space. In *Stochastic analysis for Poisson point processes*, pages 185–228. Springer, 2016.

[12] P. Boyle and T. Wang. Pricing of new securities in an incomplete market: the catch 22 of no-arbitrage pricing. *Mathematical Finance*, 11(3):267–284, 2001.

[13] L. H. Y. Chen. On the convergence of Poisson binomial to Poisson distributions. *Ann. Probability*, 2(1):178–180, 1974.

[14] H. Chernoff. A note on an inequality involving the Normal distribution. *The Annals of Probability*, pages 533–535, 1981.

[15] J. Cox, S. Ross, A, and M. Rubinstein. Option pricing: A simplified approach. *Journal of financial Economics*, 7(3):229–263, 1979.

[16] R. Dalang, A. Morton, and W. Willinger. Equivalent martingale measures and no-arbitrage in stochastic securities market models. *Stochastics: An International Journal of Probability and Stochastic Processes*, 29(2):185–201, 1990.

[17] D. J. Daley and D. Vere-Jones. *An introduction to the theory of point processes: volume II: general theory and structure*. Springer Science & Business Media, 2007.

[18] L. Decreusefond. The Stein-Dirichlet-Malliavin method. *ESAIM: Proc.*, 51:49–59, 2015.

[19] L. Decreusefond and I. Flint. Moment formulae for general point processes. *Comptes Rendus Mathematique*, 352(4):357 – 361, 2014.

[20] L. Decreusefond and H. Halconruy. Malliavin and Dirichlet structures for independent random variables. *Stochastic Processes and their Applications*, 129(8):2611–2653, Aug 2019.

[21] L. Decreusefond and P. Moyal. *Stochastic modeling and analysis of telecom networks*. John Wiley & Sons, 2012.

[22] F. Delbaen and W. Schachermayer. *The Mathematics of Arbitrage*. Springer-Verlag Berlin Heidelberg, 2006.

[23] C. Döbler and G. Peccati. The fourth moment theorem on the Poisson space. *Ann. Probab.*, 46(4):1878–1916, 07 2018.

[24] M. Duerinckx, A. Gloria, and F. Otto. The structure of fluctuations in stochastic homogenization. *Communications in Mathematical Physics*, pages 1–48, 2020.

[25] T. Erhardsson. Stein's method for Poisson and compound Poisson. *An introduction to Stein's method*, 4.

[26] H. Föllmer and M. Schweizer. Hedging of contingent claims under incomplete information. *Applied stochastic analysis*, 5(389-414):19–31, 1991.

[27] H. Föllmer and D. Sondermann. Hedging of non-redundant contingent claims. 1985.

[28] M. Frittelli. The minimal entropy martingale measure and the valuation problem in incomplete markets. *Mathematical finance*, 10(1):39–52, 2000.

[29] M. Geske, A. Godbole, A. Schaffner, A. Skolnick, and G. Wallstrom. Compound Poisson approximations for word patterns under markovian hypotheses. *Journal of applied probability*, pages 877–892, 1995.

[30] N. Hakansson. The fantastic world of finance: Progress and the free lunch. *Journal of Financial and Quantitative Analysis*, pages 717–734, 1979.

[31] H. Halconruy. *Calcul de Malliavin et structures de Dirichlet pour des variables aléatoires indépendantes*. PhD thesis, Institut polytechnique de Paris, 2020.

[32] J. Jacod and A. Shiryaev. Local martingales and the fundamental asset pricing theorems in the discrete-time case. *Finance and stochastics*, 2(3):259–273, 1998.

[33] K. Krokowski. Poisson approximation of Rademacher functionals by the Chen-Stein method and Malliavin calculus. *arXiv preprint arXiv:1505.01417*, 2015.

[34] R. Lachièze-Rey and G. Peccati. Fine Gaussian fluctuations on the Poisson space II: rescaled kernels, marked processes and geometric U-statistics. *Stochastic Processes and their Applications*, 123(12):4186–4218, December 2013.

[35] G. Last. Stochastic analysis for Poisson processes. In *Stochastic analysis for Poisson point processes*, pages 1–36. Springer, 2016.

[36] G. Last, G. Peccati, and M. Schulte. Normal approximation on Poisson spaces: Mehler's formula, second order Poincaré inequalities and stabilization. August 2016.

[37] G. Last and M. Penrose. Martingale representation for Poisson processes with applications to minimal variance hedging. *Stochastic Processes and their Applications*, 121(7):1588 – 1606, 2011.

[38] G. Last and M. Penrose. *Lectures on the Poisson Process.* Institute of Mathematical Statistics Textbooks. Cambridge University Press, 2017.

[39] M. Ledoux. Chaos of a Markov operator and the fourth moment condition. *The Annals of Probability*, 40(6):2439–2459, 2012.

[40] J Michael M. Harrison and D. Kreps. Martingales and arbitrage in multiperiod securities markets. *Journal of Economic theory*, 20(3):381–408, 1979.

[41] M. Métivier. Semimartingales, volume 2 of de gruyter studies in mathematics, 1982.

[42] J. Nash. Continuity of solutions of parabolic and elliptic equations. *American Journal of Mathematics*, 80(4):931–954, 1958.

[43] I. Nourdin and G. Peccati. Stein's method on Wiener chaos. *Probability Theory and Related Fields*, 145(1-2):75–118, June 2008.

[44] I. Nourdin and G. Peccati. Noncentral convergence of multiple integrals. *The Annals of Probability*, 37(4):1412–1426, 2009.

[45] I. Nourdin and G. Peccati. *Normal Approximations with Malliavin Calculus: From Stein's Method to Universality.* Cambridge University Press, 2012.

[46] I. Nourdin, G. Peccati, and G. Reinert. Invariance principles for homogeneous sums: universality of Gaussian Wiener chaos. *The Annals of Probability*, 38(5):1947–1985, 2010.

[47] I. Nourdin and G Peccati, G.and Reinert. Stein's method and stochastic analysis of Rademacher functionals. *Electronic Journal of Probability*, 15:1703–1742, 2010.

[48] D. Nualart. *The Malliavin Calculus and Related Topics.* Springer-Verlag, 1995.

[49] D. Nualart. Normal approximation on a finite Wiener chaos. In Dan Crisan, Ben Hambly, and Thaleia Zariphopoulou, editors, *Stochastic Analysis and Applications 2014*, pages 377–395, Cham, 2014. Springer International Publishing.

[50] D. Nualart and W. Schoutens. Chaotic and predictable representations for Lévy processes. *Stochastic Processes and their Applications*, 90(1):109 – 122, 2000.

[51] D. Nualart and J. Vives. Anticipative calculus for the Poisson process based on the Fock space. In *Séminaire de probabilités XXIV*, pages 154–165. Springer-Verlag, 1988.

[52] G. Di Nunno, B. Oksendal, and F. Proske. White noise analysis for Lévy processes. *Journal of Functional Analysis*, 206:109–148, 2004.

[53] G. Peccati. The Chen-Stein method for Poisson functionals. *arXiv preprint arXiv:1112.5051*, 2011.

[54] G. Peccati, J. L. Solé, M. S. Taqqu, and F. Utzet. Stein's method and Normal approximation of Poisson functionals. *Ann. Probab.*, 38(2):443–478, 03 2010.

[55] N. Privault. Chaotic and variational calculus in discrete and continuous time for the Poisson process. *Stoch. & Stoch. Rep.*, 51:83–109, 1994.

[56] N. Privault. *Stochastic Analysis in Discrete and Continuous Settings*. 1982. Springer-Verlag Berlin Heidelberg, 2009.

[57] N. Privault. *Stochastic finance: An introduction with market examples*. CRC Press, 2013.

[58] N. Privault. Chapter 19 Stochastic calculus for jump processes. 2018.

[59] N. Privault and G. Torrisi. The Stein and Chen-Stein methods for functionals of non-symmetric Bernoulli processes. *ALEA Lat. Am. J. Probab. Math. Stat*, 12(1):309–356, 2015.

[60] G. Reinert and S. Schbath. Compound Poisson and Poisson process approximations for occurrences of multiple words in markov chains. *Journal of Computational Biology*, 5(2):223–253, 1998.

[61] R. Rendleman and B. Bartter. Two-state option pricing. *The Journal of Finance*, 34(5):1093–1110, 1979.

[62] S. Robin. A compound Poisson model for word occurrences in dna sequences. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 51(4):437–451, 2002.

[63] S. Robin and S. Schbath. Numerical comparison of several approximations of the word count distribution in random sequences. *Journal of Computational Biology*, 8(4):349–359, 2001.

[64] W. Runggaldier. *Portfolio optimization in discrete time*. Accademia delle Scienze dell'Istituto di Bologna, 2006.

[65] W. Schachermayer. A Hilbert space proof of the fundamental theorem of asset pricing in finite discrete time. *Insurance: Mathematics and Economics*, 11(4):249–257, 1992.

[66] S. Schbath. Compound poisson approximation of word counts in DNA sequences. *ESAIM: probability and statistics*, 1:1–16, 1997.

[67] M. Schulte. Normal approximation of Poisson functionals in Kolmogorov distance. *Journal of theoretical probability*, 29(1):96–117, 2016.

[68] M. Schweizer. Option hedging for semimartingales. *Stochastic Processes and their Applications*, 37(2):339 – 363, 1991.

[69] M. Schweizer. Mean-variance hedging for general claims. *The annals of applied probability*, pages 171–179, 1992.

[70] M. Schweizer. Variance-optimal hedging in discrete time. *Mathematics of Operations Research*, 20(1):1–32, 1995.

[71] C. Stein. A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, Volume 2: Probability Theory*, pages 583–602, Berkeley, Calif., 1972. University of California Press.

[72] G. Torrisi. Poisson approximation of point processes with stochastic intensity, and application to nonlinear Hawkes processes. In *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, volume 53, pages 679–700. Institut Henri Poincaré, 2017.

[73] J. J. Víquez R. Normal convergence using Malliavin calculus with applications and examples. *Stochastic Analysis and Applications*, 36(2):341–372, 2018.

[74] L. Wu. A new modified logarithmic Sobolev inequality for Poisson point processes and several applications. *Probability Theory and Related Fields*, 118(3):427–438, 2000.

[75] G. Zheng. Normal approximation and almost sure central limit theorem for non-symmetric rademacher functionals. 2017.