Electronic Journal of Probability

Electron. J. Probab. **27** (2022), article no. 164, 1–39.
ISSN: 1083-6489 https://doi.org/10.1214/22-EJP892

Malliavin calculus for marked binomial processes and applications*

Hélène Halconruy[‡]

Abstract

We develop stochastic analysis tools for marked binomial processes (MBP) that are the discrete analogues of the marked Poisson processes. They include in particular: (i) the statement of a chaos decomposition for square-integrable functionals of MBP, (ii) the design of a tailor-made Malliavin calculus of variations, (iii) the statement of the analogues of Stroock's, Clark's and Mehler's formulas. We provide our formalism with two applications: (App1) studying the (compound) Poisson approximation of MBP functional by combining it with the Chen-Stein method and (App2) solving an optimal hedging problem in the trinomial model.

Keywords: Malliavin calculus; chaos expansion; Chen-Stein method; Poisson limit theorems; trinomial market model; optimal hedging.
MSC2020 subject classifications: Primary 60H07; 60J75; 60G55, Secondary 60F05; 91G10.
Submitted to EJP on April 1, 2021, final version accepted on December 12, 2022.
Supersedes arXiv:2104.00914.

1 Introduction

The aim of this paper is to develop a stochastic analysis for marked binomial processes (MBP); these are the discrete analogues of the marked Poisson processes, i.e., point processes that are defined on $\mathbb{N} \times \mathrm{E}$, where $\mathbb{N} := \{n,\, n \geqslant 1\}$ and E is a state space of random marks.
Our main theoretical achievements (described in section 3) are:

(**I**) the statement of a chaos decomposition for square-integrable functionals of MBP, in terms of multiple discrete stochastic integrals with respect to some specific normal martingales (Theorem 3.8);

*This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N°811017.
[†]Léonard de Vinci Pôle universitaire, Research Center, 92 916 Paris La Défense, France. E-mail: helene.halconruy@devinci.fr
[‡]Mathematic research unit, Université du Luxembourg, Esch-sur-Alzette, Luxembourg

(II) the design of a tailor-made Malliavin calculus of variations, from suitable versions of classical operators (the Malliavin derivative, the divergence operator, the Ornstein-Uhlenbeck generator and its pseudo-inverse), and an integration by parts formula (Proposition 3.15);

(III) the statement of the discrete analogues of three useful functional identities already established in the Gaussian and the Poisson cases: Stroock's formula (Theorem 4.1), Clark's formula (Theorem 4.4) and Mehler's formula (Theorem 4.9).

Our findings represent natural, yet powerful counterparts to existing results on the Gaussian (see e.g. [37, 38]) and Poisson (see [10] for a variational approach, [40, 45] for a chaotic approach) spaces. They also find their place among discrete-time theories: comparable calculus have been developed on the Rademacher space in [46], for i.i.d. random variables in [11, 15, 19, 48], and for obtuse random walks in [25]. In particular, our analysis unfolds from the decomposition in chaos of square-integrable functionals and revolves around *ad-hoc* versions of classical operators like the Malliavin derivative, the divergence operator, the Ornstein-Uhlenbeck generator and its pseudo-inverse. As by-products of this formalism, we can state three discrete-time functional identities that can be understood compared to their Poisson counterparts: Stroock's formula stated in [32], Clark's formula in [32, 58] and Mehler's formula in [31]. They provide applications that justify our study.

Indeed, as described in sections 5 and 6, we stress that our main motivations come from two distinct applications.

(App 1) The implementation of the *Malliavin-Stein method*; introduced in [36, 37] to deal with Normal approximation on a Gaussian space, we use it here to assess the total variation distance between the law of a (compound) Poisson random variable and that of a (sufficiently regular) integer-valued MBP functional. Stein's method, initially developed to quantify the rate of convergence in the Central Limit Theorem in [55] was further adapted to deal with the Poisson approximation of sums of dependent variables in [12]. The so-called Chen-Stein method has been first performed to quantify the Poisson approximation [1, 2, 8] or the compound Poisson one [6, 9] by sums of possibly dependent indicators random variables. It relies on location arguments around the notion of *neighbourhood of dependence* sets. In the same spirit of [43], we combine Chen-Stein method with the Malliavin calculus to quantify the (compound) Poisson approximation of integer-valued MBP functionals and revisit two related issues: the Poisson approximation of the length of the longest head run in a series of independent coin tosses in Subsection 5.1 and the compound Poisson approximation of the number of occurrences of a rare word in a DNA sequence in Subsection 5.2.

(App 2) The explicit computation of the squared-loss minimizing strategy in a trinomial-type model in Section 6. In some *complete* market models, the replication strategy writes in terms of Malliavin derivative through the Clark-Ocone formula: historically highlighted in the Black-Scholes model as in [42, 28], this has also been shown in the Cox-Ross-Rubinstein (chapter one in [46] and in [47]) or in more generalized discrete-time complete market models that can be ingeniously built from an obtuse random walk as in [25]. In *incomplete* markets like the trinomial model, not all claims are attainable so that we are led to consider a loss minimizing hedging problem as in [24, 52, 54]. By replacing the trinomial model by a surrogate (equivalent) one driven by a MBP that we call *jump-binomial model*, and combining our tools with the minimal variance approach (see [54]), we give an explicit expression of the squared-loss minimizing strategy in the trinomial/jump-binomial model in terms of the Malliavin derivative for MBP.

The paper is structured as follows. Section 2 is devoted to the framework and main notations.

Notation The following set notations are used throughout the paper. Let us denote $\mathbb{N} = \{n, n \geqslant 1\}$ and $\mathbb{Z}_+ = \mathbb{N} \cup \{0\}$. We consider a measurable space $(\mathbb{X}, \mathcal{X})$ where $\mathbb{X} = \mathbb{N} \times \mathrm{E}$, where E is a countable subset of $\mathbb{Z}$. We denote $[\![a, b]\!] := \{a, \dots, b\}$ for any $a, b \in \mathbb{Z}$ such that $a < b$; in the particular case where $a = 1$, we will opt for the lighter notation $[b] := \{1, \dots, b\}$. By convention, $[\![1, 0]\!] = \emptyset$. For any $t \in \mathbb{N}$ let $\mathbb{X}_t := [t] \times \mathrm{E}$. Any n-tuple of $\mathbb{X}^n$ can be denoted by bold letters; for instance, $(\mathbf{t}_n, \mathbf{k}_n) = \big((t_1, k_1), \dots, (t_n, k_n)\big)$. For any $\mathrm{A} \in \mathcal{X}$, we denote $\mathrm{A}^{n,<} = \{(\mathbf{t}_n, \mathbf{k}_n) \in \mathrm{A}^n \,:\, t_1 < \cdots < t_n\}$, 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.

2 Framework and main notation

Setup $(\Omega, \mathcal{A}, \mathbf{P})$ will hereafter be an abstract probability space assumed to be wide enough to support all random objects in question.

The mark space Consider the measurable space $(\mathbb{X}, \mathcal{X})$ where $\mathbb{X} = \mathbb{N} \times \mathrm{E}$ and E is a countable (possibly finite) subset of $\mathbb{Z}$ that we call the *mark set*. Denote by $\mathfrak{N}_{\mathbb{X}}$ (resp. $\widehat{\mathfrak{N}}_{\mathbb{X}}$) the space of simple, integer-valued, σ-finite (resp. finite) measures on $\mathbb{X}$.

Let $\mathcal{N}^{\mathbb{X}}$ be the smallest σ-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 $\mathrm{A} \in \mathcal{X}$.

Point processes A point process (resp. finite point process) is a random element η in $\mathfrak{N}_{\mathbb{X}}$ (resp. in $\widehat{\mathfrak{N}}_{\mathbb{X}}$) that satisfies $\eta(\mathrm{A}) \in \mathbb{Z}_+ \cup \{\infty\}$ (resp. $\eta(\mathrm{A}) \in \mathbb{Z}_+$) for all $\mathrm{A} \in \mathcal{X}$. In this simple frame, we may and will assume that any element η of $\mathfrak{N}_{\mathbb{X}}$ is *proper*, i.e., it can be 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 $x \in \mathbb{X}$, δ_x is the Dirac measure at x. For a complete exposé on the subject of point processes, the reader can refer to the section 6.1 of the monograph [33], or to [30].

Marked binomial process Let $\lambda \in (0, 1)$. The underlying marked process η is the random element of $(\mathfrak{N}_{\mathbb{X}}, \mathcal{N}^{\mathbb{X}})$ that is a measurable map from $(\Omega, \mathcal{A})$ to $(\mathfrak{N}_{\mathbb{X}}, \mathcal{N}^{\mathbb{X}})$. In analogy with marked Poisson processes (see [33], chapter 7), we can represent the marked binomial process as an element of $\mathfrak{N}_{\mathbb{X}}$ such that for all $\mathrm{A} \in \mathcal{X}$ and $\mathbf{P}$-almost surely all $\omega \in \Omega$,

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

where:

- $(\mathrm{T}_t)_{t \geqslant 0}$ is a sequence of jump times such that $\mathrm{T}_0 = 0$, $\mathrm{T}_t := \sum_{s=1}^{t} \xi_s$ ($t \in \mathbb{N}$) and the inter-arrival random variables $\{\xi_t,\, t \in \mathbb{N}\}$ are i.i.d. geometric random variables supported on $\mathbb{N}$ with mean $1/\lambda$, i.e., such that $\mathbf{P}(\{\xi_t = k\}) = \lambda(1 - \lambda)^{k-1}$ for $k \geqslant 1$.
- $\{\mathrm{V}_t,\, t \in \mathbb{N}\}$ is a collection of i.i.d. E-valued random elements such that $\mathbf{P}$-almost surely $\eta(\mathrm{T}_t, \mathrm{V}_t) = 1$ for $\mathrm{T}_t < \infty$, and that is assumed to be independent of the T_t.

By a slight abuse of notation, we shall use the shortcut $\eta(t, k) = \eta(\{(t, k)\})$ for $(t, k) \in \mathbb{X}$. We can construct the marked binomial process as follows:

1. Consider an infinite number of independent Bernoulli experiments where a *success* stands for a jump and occurs with probability λ. The random variable N_t that counts the number of jumps (or successes) until time t follows a binomial distribution $\mathcal{B}\mathrm{in}(t, \lambda)$.

2. If there is a jump at time t, i.e., if there exists $s \in \mathbb{N}$ such that $\mathrm{T}_s = t$, draw a mark $k \in \mathrm{E}$ according to the distribution $\mathbf{Q}$ and let $\eta(t,k) = 1$. Otherwise, if there is no jump at time t, let $\eta(t,k) = 0$ for all $k \in \mathrm{E}$ and $\eta(t,\cdot) = 0$ where we define $\eta(t,\cdot)$ by

$$\eta(t,\cdot) := \sum_{k \in \mathrm{E}} \eta(t,k).$$

The underlying jump process $\mathrm{N} = (\mathrm{N}_t)_{t \geqslant 0}$ is defined by $\mathrm{N}_0 = 0$ and $\mathrm{N}_t = \sum_{s \in \mathbb{N}} \mathbf{1}_{\{\mathrm{T}_s \leqslant t\}}$. Note that for any $t \in \mathbb{N}$, N_t is a binomial random variable with mean λt. The random variables $\Delta \mathrm{N}_t := \mathrm{N}_t - \mathrm{N}_{t-1} = \eta(t,\cdot)$ are independent Bernoulli random variables, and for any $k, \ell \in \mathrm{E}$, $t, s \in \mathbb{N}$ such that $t \neq s$, **the random variables** $\eta(t,k)$ **and** $\eta(s,\ell)$ **are independent**.

Last, let $\mathrm{W}_t = \mathrm{V}_{\mathrm{N}_t}$ be the random variable indicating the mark associated with a jump occurring at time t. In fact, η can be viewed as the time-ordered sequence $((\mathrm{T}_t, \mathrm{V}_t))_{t \geqslant 1}$. This means that η can be identified to the set of elements of $\mathbb{X}$ it weights as illustrated below.

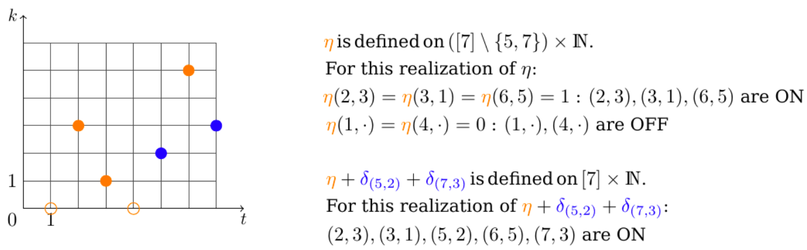

Figure 1: **Realization of a MBP on** $\mathbb{X} = [7] \times \mathbb{N}$

In the case $\mathrm{E} = \{1\}$, $\mathbb{X}$ can be identified to $\mathbb{N}$ and we refer to η as a *simple binomial process*.

Associated filtered probability space For any $\mathrm{A} \in \mathcal{X}$, $\eta(\mathrm{A})$ is the map $\omega \in \Omega \mapsto \eta(\mathrm{A})(\omega)$. We may (and will) assume that $\mathcal{A} = \sigma(\{\eta(\mathrm{A}) = k,\ k \geqslant 0\})$ and $\mathcal{F} := (\mathcal{F}_t)_{t \geqslant 0}$ is the canonical filtration defined from η by

$$\mathcal{F}_0 := \{\emptyset, \Omega\} \quad \text{and} \quad \mathcal{F}_t := \sigma\big(\eta(s,k),\ s \leqslant t, k \in \mathrm{E}\big),\ t \geqslant 1.$$

Let $\mathbf{Q}$ be the common distribution of the V_t and $\mathbf{P}_\eta = \mathbf{P} \circ \eta^{-1}$ be the image measure of $\mathbf{P}$ under η on the space $(\mathfrak{N}_{\mathbb{X}}, \mathcal{N}^{\mathbb{X}})$, i.e., the distribution of η. Let us define ν the measure on $\mathbb{X}$ such that

$$\nu(t,k) := \mathbf{E}\left[\sum_{s=1}^{\mathrm{N}_t} \mathbf{1}_{\{\mathrm{V}_s = k\}}\right] - \mathbf{E}\left[\sum_{s=1}^{\mathrm{N}_{t-1}} \mathbf{1}_{\{\mathrm{V}_s = k\}}\right] = \lambda t p_k - \lambda(t-1)p_k = \lambda p_k =: \lambda \mathbf{Q}(\{k\}),$$

where the second equality comes from Wald's lemma. The intensity of η – viewed as a random measure on $\mathbb{X}$ – is defined for any $\mathrm{A} \in \mathcal{X}$ by

$$\nu(\mathrm{A}) = \mathbf{E}[\eta(\mathrm{A})] = \sum_{(t,k) \in \mathrm{A}} \nu(t,k). \tag{2.3}$$

Random variables on $\mathfrak{N}_{\mathbb{X}}$ We denote hereafter by $\mathbb{R}(\mathfrak{N}_{\mathbb{X}})$ 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 Ω. Since $\mathcal{A} = \sigma(\eta)$, for any $\mathrm{F} \in \mathcal{L}^0(\Omega)$, there exists a function $\mathfrak{f} \in \mathbb{R}(\mathfrak{N}_{\mathbb{X}})$ 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 denoted by the corresponding Gothic lowercase letter, $\mathfrak{f}$. Last, for $p \in \mathbb{N}$, we define $\mathrm{L}^p(\mathbf{P}) := \mathrm{L}^p(\Omega, \mathcal{A}, \mathbf{P})$ the set of p-integrable functions on Ω with respect to $\mathbf{P}$.

The *compound binomial process* $\mathrm{Y} = (\mathrm{Y}_t)_{t \geqslant 1}$ of intensity ν is defined by

$$\mathrm{Y}_t := \sum_{s=1}^{\mathrm{N}_t} \mathrm{V}_s = \sum_{s=1}^{t} \Delta \mathrm{N}_s \mathrm{W}_s. \tag{2.4}$$

By Wald's lemma, we can check that

$$\mathbf{E}\big[\mathrm{Y}_t - \mathrm{Y}_{t-1} | \mathcal{F}_{t-1}\big] = \sum_{k \in \mathrm{E}} \lambda k p_k = \lambda \mathbf{E}[\mathrm{V}_1].$$

The corresponding *compensated compound binomial process*, i.e., the $(\mathbf{P}, \mathcal{F})$-martingale $\overline{\mathrm{Y}} := (\overline{\mathrm{Y}}_t)_{t \in \mathbb{T}}$, is defined by $\overline{\mathrm{Y}}_0 = 0$ and for any $t \in \mathbb{N}$,

$$\overline{\mathrm{Y}}_t = \Big(\sum_{s=1}^{\mathrm{N}_t} \mathrm{V}_s \Big) - \sum_{(s,k) \in \mathbb{X}_t} \lambda k p_k = \sum_{(s,k) \in \mathbb{X}_t} k \Big[\mathbf{1}_{\{\eta(s,k)=1\}} - \lambda p_k \Big] =: \sum_{(s,k) \in \mathbb{X}_t} k \Delta \mathrm{Z}_{(s,k)}. \tag{2.5}$$

The elements of the family $\mathcal{Z} := \{\Delta \mathrm{Z}_{(t,k)},\ (t,k) \in \mathbb{X}\}$ are such that: (i) $\mathbf{E}[\Delta \mathrm{Z}_{(t,k)}] = 0$, (ii) $\Delta \mathrm{Z}_{(t,\cdot)}$ and $\Delta \mathrm{Z}_{(s,\cdot)}$ are independent for $s \neq t$, (iii) For all $t \geqslant 1$, the random variables $\Delta \mathrm{Z}_{(t,k)}$ and $\Delta \mathrm{Z}_{(1,k)}$ are identically distributed and centered, since we have

$$\mathbf{E}\big[\mathbf{1}_{\{\eta(t,k)=1\}}\big] = \nu(t,k) = \lambda \mathbf{Q}(\{k\}) = \mathbf{E}\big[\mathbf{1}_{\{\eta(1,k)=1\}}\big], \tag{2.6}$$

by definition (2.3) of ν. However for $t \in \mathbb{N}$, $k, \ell \in \mathrm{E}$, $\Delta \mathrm{Z}_{(t,k)}$ and $\Delta \mathrm{Z}_{(t,\ell)}$ are not orthogonal for the inner product $(\mathrm{X}, \mathrm{Y}) \mapsto \mathbf{E}[\mathrm{XY}]$.

3 Main theoretical results

In this section we develop stochastic analysis tools for marked binomial processes (MBP). In particular: (i) we establish a chaos decomposition for any square-integrable functional of MBP, (ii) we define a *modicum* of Malliavin operators (gradient, divergence, Ornstein-Uhlenbeck operators), (iii) we connect these operators to their analogues in L^1.

3.1 Chaos decomposition

To define marked binomial chaoses, we build a discrete analogue of the Itô-Wiener integral.

3.1.1 (Multiple) stochastic integrals

Processes 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 $u = \sum_{(t,k) \in \mathbb{X}} \mathfrak{u}(\eta, (t,k)) \mathbf{1}_{\{\eta(t,k)=1\}}$, 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 $\mathbb{R}$ and $\mathfrak{u}$ is called the *representative* of u. As for random variables, the representative of a process will be denoted by a Gothic letter.

Definition 3.1. *The set of simple processes, denoted by* $\mathcal{U}$*, is the set of random variables of the form*

$$u = \sum_{(t,k) \in \mathbb{X}_T} \mathfrak{u}(\eta, (t,k)) \mathbf{1}_{\{(t,k)\}}, \tag{3.1}$$

where $T \in \mathbb{N}$*, and* $\mathfrak{u}$ *is the representative of* u*. Let* $\mathcal{P}$ *be the subspace of* $\mathcal{U}$ *made of simple predictable processes, i.e., of the form* (3.1) *where* $\mathfrak{u}(\eta, (t, \cdot))$ *is* $\mathcal{F}_{t-1}$*-measurable for any* $t \in [T]$*.*

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

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

and

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

Normal martingales We define the analogue of the Itô-Wiener integral for any function that is square-integrable with respect to the intensity measure. This requires that the integral satisfies an isometry property; since the family $\mathcal{Z}$ is not orthogonal, we can not construct the Itô-Wiener integral – as usual – from the increments of the process $\overline{\mathrm{Y}}$ defined by (2.5). Let us then define a suitable random variable family for integration, linked to *normal martingales*. For $T\in\mathbb{N}$, define $\mathcal{Z}_T=\{\Delta\mathrm{Z}_{(t,k)}\,;\,(t,k)\in\mathbb{X}_T\}$. The dimension of the related spanned space is equal to

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

so that we can derive from $\mathcal{Z}_T$ an orthogonal family, $\mathcal{R}_T=\{\Delta\mathrm{R}_{(t,k)}\,;\,(t,k)\in\mathbb{X}_T\}$. Assume that $\mathrm{E}=\{k^1,\dots,k^{\overline{\mathfrak{m}}}\}$; then, the Gram-Schmidt process provides $\Delta\mathrm{R}_0=1$,

$$\Delta\mathrm{R}_{(t,k^1)}=\Delta\mathrm{Z}_{(t,k^1)}\text{ and }\ \Delta\mathrm{R}_{(t,k^n)}=\Delta\mathrm{Z}_{(t,k^n)}-\sum_{j=1}^{n-1}\frac{\mathbf{E}\big[\Delta\mathrm{Z}_{(1,k^n)}\Delta\mathrm{R}_{(1,k^j)}\big]}{\mathbf{E}\big[(\Delta\mathrm{R}_{(1,k^j)})^2\big]}\Delta\mathrm{R}_{(t,k^j)},\tag{3.2}$$

for $n\in[\overline{\mathfrak{m}}]$, by noting that (i) the random variables $\Delta\mathrm{R}_{(t,k)}$ and $\Delta\mathrm{R}_{(1,k)}$ are identically distributed since $\Delta\mathrm{Z}_{(t,k)}$ and $\Delta\mathrm{Z}_{(1,k)}$ are (as a consequence of (2.6)) and (ii) that for any $s\leqslant t-1$, $\mathbf{E}\big[\Delta\mathrm{R}_{(s,k)}\,\Delta\mathrm{Z}_{(t,\ell)}\big]=\mathbf{E}\Big[\Delta\mathrm{R}_{(s,k)}\mathbf{E}\big[\Delta\mathrm{Z}_{(t,\ell)}|\mathcal{F}_s\big]\Big]=0$, for $k,\ell\in\mathrm{E}$.

For any $t\in[T]$, $(\Delta\mathrm{Z}_{(t,k)},\,k\in\mathrm{E})$ is actually the image of $(\Delta\mathrm{R}_{(t,k)},\ k\in\mathrm{E})$ by the linear transformation associated to the $\overline{\mathfrak{m}}\times\overline{\mathfrak{m}}$ triangular matrix $\mathfrak{M}$ defined by

$$\mathfrak{M}=(\mathfrak{m}_{ij})_{i,j\in[\![1,\overline{\mathfrak{m}}]\!]}=\begin{pmatrix}1&0&\cdots&0\\ \gamma_{21}&1&\cdots&0\\ \vdots&\vdots&\ddots&\vdots\\ \gamma_{n1}&\gamma_{n2}&\cdots&1\end{pmatrix},\tag{3.3}$$

where $\gamma_{ij}:=\mathbf{E}[\Delta\mathrm{Z}_{(1,k^i)}\Delta\mathrm{R}_{(1,k^j)}]/\mathbf{E}[(\Delta\mathrm{R}_{(1,k^j)})^2]$ for $i<j$. As $\mathfrak{M}$ is invertible, $(\Delta\mathrm{R}_{(t,k)},\,k\in\mathrm{E})$ is obtained via the product of $\mathfrak{M}^{-1}$ by the vector $(\Delta\mathrm{Z}_{(t,k)},\ k\in\mathrm{E})$. Since this linear transformation is bijective, the family $\mathcal{R}$ can be constructed in a similar fashion when E is countably infinite, by induction.

Remark 3.2. (1) We can construct such a family $\mathcal{R}$ even if E is not countable (if for instance $\mathrm{E}=\mathbb{R}$), by drawing inspiration from the *orthogonal power jump process* for Lévy processes. This was introduced in [39] to define the so-called *Teugels integrals* and taken over in [41] to state a generalization of the Clark-Ocone formula for Lévy processes. Transposing it into our framework, that would give: Define for any $n\in\mathbb{N}$,

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

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 γ_{nj} are real numbers such that the processes of the collection $\{(\Delta\mathrm{R}_t^{(n)})_{t\geqslant1},\ n\in\mathbb{N}\}$ are *strongly* *orthogonal martingales*, i.e., for any $t \in \mathbb{N}$, the product $\Delta \mathrm{R}^{(n)} \Delta \mathrm{R}^{(m)}$ is a uniformly integrable martingale for all $(n, m) \in \mathbb{N}^2$, $m \neq n$.

(2) In the different frame of *obtuse random walks*, a construction of the discrete-time analogues of *normal martingales* in $\mathbb{R}^n$ (resp. in $\mathbb{C}^n$) from an initial process that satisfies the *predictable representation theorem* was proposed in [4] (resp. in [3]). These tools were reused in [25] to state a Clark-Ocone-type formula and were further adapted to deal with discrete-time processes on the sequence space in [22].

Integrals of square-integrable predictable processes These are defined as the discrete stochastic integrals with respect to the family $\mathcal{R}$. Let $\kappa_k := \mathbf{E}[\Delta \mathrm{R}^2_{(t,k)}]$ and $\tilde{\nu}$ be the measure on $\mathbb{X}$ such that for $(t, k) \in \mathbb{X}$,

$$\tilde{\nu}(t, k) = \kappa_k \nu(t, k). \tag{3.4}$$

Proposition 3.3. *Any $u \in \mathcal{U}$ of representative $\mathfrak{u}$ is integrable with respect to the family $\mathcal{R}$ by*

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

The integral J_1 extends to square-integrable predictable processes via the isometry formula

$$\mathbf{E}\Big[\big|\mathrm{J}_1(u)\big|^2\Big] = \mathbf{E}\Big[\|u\|^2_{\mathrm{L}^2(\mathbb{X}, \tilde{\nu})}\Big], \tag{3.5}$$

where $\tilde{\nu}$ is the measure on $\mathbb{X}$ defined by $\tilde{\nu}(t, k) = \kappa_k \nu(t, k)$ and $\kappa_k := \mathbf{E}[\Delta \mathrm{R}^2_{(t,k)}]$ for $(t, k) \in \mathbb{X}$.

Proof. The proof is close to that of Proposition 1.3.2. in [46] and won't be detailed. It relies on the fact that $\Delta \mathrm{R}_{(t,k)}$ and $\mathcal{F}_{t-1}$ are independent, $\Delta \mathrm{R}_{(t,k)}$ is centered and $\mathbf{E}[\Delta \mathrm{R}^2_{(t,k)}] = \kappa_k$. ☐

Discrete Itô-Wiener type integral For $f \in \mathrm{L}^2(\mathbb{X}, \nu)$, the discrete type integral denoted by $\mathrm{J}_1(f)$ is defined by

$$\mathrm{J}_1(f) = \sum_{(t,k) \in \mathbb{X}} f(t, k) \Delta \mathrm{R}_{(t,k)}.$$

Let us also define for $f \in \mathrm{L}^2(\mathbb{X}, \nu)$,

$$\mathrm{J}_1(f\,;\,\mathcal{Z}) = \sum_{(t,k) \in \mathbb{X}} f(t, k) \Delta \mathrm{Z}_{(t,k)}. \tag{3.6}$$

To define multiple stochastic integrals, we work in a space of symmetric functions. The space $\mathrm{L}^2(\mathbb{X}, \nu)^{\circ 0}$ where ν is defined by (2.3) is by convention identified to $\mathbb{R}$; let thus for any $f \in \mathrm{L}^2(\mathbb{X}, \nu)^{\circ 0}$, $\mathrm{J}_0(f_0) = f_0$.

Definition 3.4. *For $n \in \mathbb{N}$, 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 $f_n \in \mathbb{R}(\mathbb{X}^n)$ symmetric in their n variables, i.e., such that for any permutation τ of $\{1, \dots, n\}$, $f_n\big((t_{\tau(1)}, k_{\tau(1)}), \dots, (t_{\tau(n)}, k_{\tau(n)})\big) = f_n\big((t_1, k_1), \dots, (t_n, k_n)\big)$, 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 with the inner 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 $\mathcal{C}_K(\mathbb{X}^n, \mathbb{R})$, the set of real functions with compact support on $\mathbb{X}^n$ and extended by isometry to $\mathrm{L}^2(\mathbb{X}, \nu)^{\circ n}$.

Proposition 3.5. *The $\mathcal{R}$-stochastic integral of order n is the mapping defined on $\mathcal{C}_K(\mathbb{X}^n, \mathbb{R})$ by*

$$\mathrm{J}_n(f_n) = n \sum_{(t,k)\in\mathbb{X}} \mathrm{J}_{n-1}(f_n(\star, (t,k)))\, \Delta \mathrm{R}_{(t,k)}, \tag{3.7}$$

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

$$\mathrm{J}_n(f_n) = n! \sum_{(\mathbf{t}_n,\mathbf{k}_n)\in\mathbb{X}^{n,<}} f_n(\mathbf{t}_n,\mathbf{k}_n) \prod_{i=1}^{n} \Delta \mathrm{R}_{(t_i,k_i)}. \tag{3.8}$$

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

$$\mathbf{E}[\mathrm{J}_n(f_n)\mathrm{J}_m(g_m)] = \mathbf{1}_{\{n\}}(m) n! \langle f_n, g_n\rangle_{\mathrm{L}^2(\mathbb{X},\tilde{\nu})^{\circ n}}, \tag{3.9}$$

so that its domain can be extended to $\mathrm{L}^2(\mathbb{X},\tilde{\nu})^{\circ n} \simeq \mathrm{L}^2(\mathbb{X},\nu)^{\circ n}$ where $\tilde{\nu}$ is defined by (3.4).

When it is unambiguous, $\mathrm{L}^2(\nu^{\circ n})$ will indifferently refer to $\mathrm{L}^2(\mathbb{X},\nu)^{\circ n}$ or $\mathrm{L}^2(\mathbb{X},\tilde{\nu})^{\circ n}$.

Proof. The proof is close to that of Proposition 1.3.2. in [46] and will not be detailed. □

3.2 Marked binomial chaoses and decomposition

This subsection is devoted to the statement of a chaos decomposition (in terms of multiple integrals) for any square-integrable *marked binomial functional*, that are random variables of the form

$$\mathrm{F} = f_0 \mathbf{1}_{\{\eta(\mathbb{X})=0\}} + \sum_{n\in\mathbb{N}} \sum_{(\mathbf{t}_n,\mathbf{k}_n)\in\mathbb{X}^n} \mathbf{1}_{\{\eta(\mathbb{X})=n\}} f_n(\mathbf{t}_n,\mathbf{k}_n) \prod_{i=1}^{n} \mathbf{1}_{\{\eta(t_i,k_i)=1\}}, \tag{3.10}$$

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 3.6. *A functional* F *is cylindrical if there exists $T \in \mathbb{N}$ such that*

$$\mathrm{F} = f_0 \mathbf{1}_{\{\eta(\mathbb{X})=0\}} + \sum_{n\in\mathbb{N}} \sum_{(\mathbf{t}_n,\mathbf{k}_n)\in\mathbb{X}_T^n} \mathbf{1}_{\{\eta(\mathbb{X})=n\}} f_n(\mathbf{t}_n,\mathbf{k}_n) \prod_{i=1}^{n} \mathbf{1}_{\{\eta(t_i,k_i)=1\}}, \tag{3.11}$$

where $\mathbb{X}_T = [T] \times \mathrm{E}$.

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

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

$\mathcal{H}_n$ is called the *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 3.7. *For any $t \in \mathbb{N}$,*

$$\mathcal{L}^0(\mathbf{P},\mathcal{F}_t) = (\mathcal{H}_0 \oplus \cdots \oplus \mathcal{H}_t) \cap \mathcal{L}^0(\mathbf{P},\mathcal{F}_t). \tag{3.12}$$

Proof. It is enough to note that for any $s,t \in \mathbb{N}$, $s \leqslant t$, the space $\mathcal{H}_s \cap \mathcal{L}^0(\mathbf{P},\mathcal{F}_t)$ is generated by the orthogonal basis

$$\{1\} \cup \left\{ \prod_{i=1}^{s} \Delta \mathrm{R}_{(t_i,k_i)},\ 1 \leqslant t_1 < \cdots < t_s \leqslant t,\ (k_1,\ldots,k_s) \in \mathrm{E}^s \right\}. \tag{3.13}$$

Indeed, any element of this basis can be written in terms of multiple integrals as

$$\prod_{i=1}^{s} \Delta \mathrm{R}_{(t_i,k_i)} = \mathrm{J}_s\Big(\mathbf{1}^{<}_{\{(t_1,k_1),\ldots,(t_s,k_s)\}}\mathbf{1}_{[t]^s}\Big) \text{ and } 1 = \mathrm{J}_0(1),$$

and any $\mathrm{F} \in \mathcal{H}_s \cap \mathcal{L}^0(\mathbf{P}, \mathcal{F}_t)$ can be written as $\mathrm{F} = \mathrm{J}_s(f_s \mathbf{1}_{[t]^s})$ with $f_s \in \mathrm{L}^2(\nu^{\circ s})$. We conclude by noting that the dimensions of $(\mathcal{H}_0 \oplus \cdots \oplus \mathcal{H}_t) \cap \mathcal{L}^0(\mathbf{P}, \mathcal{F}_t)$ and $\mathcal{L}^0(\mathbf{P}, \mathcal{F}_t)$ in (3.12) 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. □

As a direct consequence of Lemma 3.7, any random variable $\mathrm{F} \in \mathcal{L}^0(\mathbf{P}, \mathcal{F}_t)$ writes

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

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

$$\mathcal{S} = \mathrm{Span}\bigg\{\bigcup_{n \geqslant 0} \mathcal{H}_n\bigg\}.$$

Its completion in $\mathrm{L}^2(\mathbf{P})$ is denoted by $\bigoplus_{n \geqslant 0} \mathcal{H}_n$. We can state the main theorem of this chapter.

Theorem 3.8. *We have the chaos decomposition*

$$\mathrm{L}^2(\mathbf{P}) = \bigoplus_{n \geqslant 0} \mathcal{H}_n. \tag{3.14}$$

In other terms, any random variable $\mathrm{F} \in \mathrm{L}^2(\mathbf{P})$ *has a unique chaos expansion given by*

$$\mathrm{F} = \mathbf{E}[\mathrm{F}] + \sum_{n \geqslant 1} \mathrm{J}_n(f_n). \tag{3.15}$$

Proof. The proof follows closely that of Proposition 1.5.3 in [46] by combining Lemma 3.7 and the denseness of $\mathcal{S}$ in $\mathrm{L}^2(\mathbf{P})$. □

We immediately deduce an expression of the covariance of two square-integrable variables in terms of their chaos decomposition.

Corollary 3.9. *For any* $\mathrm{F}, \mathrm{G} \in \mathrm{L}^2(\mathbf{P})$*,*

$$\mathrm{cov}(\mathrm{F}, \mathrm{G}) = \sum_{n \geqslant 1} n! \langle f_n, g_n \rangle_{\mathrm{L}^2(\mathbb{X}, \tilde{\nu})^{\otimes n}}.$$

Proof. Immediately using (3.15) together with Proposition 3.3. □

3.3 Doléans exponentials

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

$$\xi(h) = \mathbf{E}[\xi(h)] + \sum_{n \geqslant 1} \frac{1}{n!} \mathrm{J}_n(h^{\otimes n}). \tag{3.16}$$

The family $(\xi_t(h))_{t\geqslant 1}$ defined by $\xi_t(h) = \xi(h\mathbf{1}_{[t]})$ can be viewed as a discrete Doléans exponential solution of the equation in differences: for any $t \in \mathbb{N}, t \geqslant 2$,

$$\xi_t(h) - \xi_{t-1}(h) = \xi_{t-1}(h) \sum_{k\in \mathrm{E}} g(t,k)\Delta \mathrm{Z}_{(t,k)} = \xi_{t-1}(h)\mathrm{J}_1(g\,;\,\mathcal{Z}),$$

where $g \in \mathrm{L}^2(\nu)$ is defined in the following theorem.

Proposition 3.10. *For any $h \in \mathrm{L}^2(\nu)$, the discrete Doléans exponential* (3.16) *writes*

$$\begin{aligned}\xi(h) &= \mathbf{E}[\xi(h)] \prod_{t\geqslant 1}\Big(1 + \sum_{k\in\mathrm{E}} h(t,k)\Delta \mathrm{R}_{(t,k)}\Big) \\ &= \mathbf{E}[\xi(h)] \prod_{t\geqslant 1}\Big(1 + \sum_{k\in\mathrm{E}} g(t,k)\big(\mathbf{1}_{\{\eta(t,k)=1\}} - \lambda \mathbf{Q}(\{k\})\big)\Big), \end{aligned} \tag{3.17}$$

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

Proof. The proof is postponed as technical results to Appendix A. □

3.4 Malliavin calculus

The Malliavin calculus is embodied by a family of operators, the *gradient* D, the *divergence operator* δ, the *Ornstein-Uhlenbeck generator* L, and its *pseudo-inverse* L^{-1}, as well as a keystone *integration by parts* formula.

3.4.1 Gradient, divergence and integration by parts

Gradient As one way to develop it, we introduce the Malliavin derivative or gradient as the *annihilation operator* acting on the space $\mathrm{L}^2(\mathbf{P})$ seen via Theorem 3.8 as a Fock space.

Definition 3.11. *Let the linear operator $\mathrm{D} : \mathcal{S} \to \mathrm{L}^2(\mathbf{P} \otimes \nu)$ be defined for $\mathrm{J}_n(f_n) \in \mathcal{H}_n$ by*

$$\mathrm{D}_{(t,k)}\mathrm{J}_n(f_n) = n\,\mathrm{J}_{n-1}\big(f_n(\star,(t,k))\mathbf{1}_{\mathbb{X}^{n-1,<,\neg t}}(\star)\big), \tag{3.18}$$

where the symbol $\star$ stands for the first $n-1$ variables (t_i,k_i) (with $i \in [n-1]$) of $f_n((t_1,k_1),\dots,\ (t_{n-1},k_{n-1}),(t,k))$ and $\in\ (\mathbb{X}^{\neg t})^{n-1,<} := \{(t_1,k_1),\dots,(t_{n-1},k_{n-1}) \in \mathbb{X}^{n-1,<}, t_i \neq t\}$.

Divergence The divergence operator is defined as the *creation operator* acting on $\mathrm{L}^2(\mathbf{P})$. Let

$$\mathcal{U} = \Bigg\{ \sum_{n\in [\![0,T]\!]} \mathrm{J}_n(f_{n+1}(\star,\cdot))\,;\, f_{n+1} \in \mathrm{L}^2(\nu^{\circ n}) \otimes \mathrm{L}^2(\nu),\, n \in [\![0,T]\!],\, T \in \mathbb{N}\Bigg\}. \tag{3.19}$$

Definition 3.12. *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 the expansion of which is of the form $\sum_{n\geqslant 0}\mathrm{J}_n(f_n(\star,\cdot))$ and satisfies*

$$\sum_{n\geqslant 0}(n+1)!\|\bar{f}_{n+1}\|^2_{\mathrm{L}^2(\nu)^{n+1}} < \infty,$$

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

$$\delta\big(\mathrm{J}_n(f_{n+1}(\star,\cdot))\big) := \mathrm{J}_{n+1}(\bar{f}_{n+1}), \tag{3.20}$$

where

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

In the frame of classical Malliavin calculus, the divergence of adapted processes coincides with the Itô-Wiener integral. We get the analogue: for any $\mathrm{A} \in \mathcal{X}$, $\delta(\mathbf{1}_{\mathrm{A}}) = \eta(\mathrm{A}) - \nu(\mathrm{A})$ holds and

$$\delta(u) = \mathrm{J}_1(u) = \sum_{(t,k)\in\mathbb{X}} \mathfrak{u}(\eta,(t,k))\Delta\mathrm{R}_{(t,k)}\ ;\ u \in \mathcal{U}. \tag{3.21}$$

This property holds for any $\mathbf{P}\otimes\nu$-square integrable process $u \in \mathcal{U}$. Let $u = \mathrm{J}_{n-1}(f_n(\star,\cdot))$ for some $f_n \in \mathrm{L}^2(\nu^{\circ n})$. The predictability of u implies that $f_n(\star,(t,k)) = g_n(\star,(t,k))\mathbf{1}_{[\![1,t-1]\!]^n}(\star)$ for some $g_n \in \mathrm{L}^2(\nu^{\circ n})$. Equation (3.21) follows by writing

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

To state gradient's closability property we need an integration by parts formula, appearing as a duality relation between D and δ. Here stands its version restricted to $\mathcal{S}\times\mathcal{U}$.

Lemma 3.13 (Integration by parts formula on $\mathcal{S}\times\mathcal{U}$)**.** *For any* $(\mathrm{F},u) \in \mathcal{S}\times\mathcal{U}$*,*

$$\mathbf{E}[\mathrm{F}\delta(u)] = \mathbf{E}\big[\langle \mathrm{DF}, u\rangle_{\mathrm{L}^2(\mathbb{X},\tilde{\nu})}\big]. \tag{3.22}$$

Proof. The proof is identical to that of Theorem 1.8.2 in [46] replacing the intensity measure ν by $\tilde{\nu}$ to benefit from the isometry property. □

Corollary 3.14 (Closability)**.** *The operator* D *is closable from* $\mathrm{L}^2(\mathbf{P})$ *to* $\mathrm{L}^2(\mathbf{P}\otimes\nu)$.

Proof. Let $(\mathrm{F}_n)_{n\geqslant 0}$ be a sequence of random variables defined on $\mathcal{S}$ such that F_n converges to 0 in $\mathrm{L}^2(\mathbf{P})$ and the sequence $(\mathrm{DF}_n)_{n\geqslant 0}$ converges to Λ in $\mathrm{L}^2(\mathbf{P}\otimes\tilde{\nu})$. Let u be a simple process of the form (3.1) for some $T \in \mathbb{N}$. From the integration by parts formula (3.22),

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

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,

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

Then,

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

for any 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 $\mathrm{L}^2(\Omega\times\mathbb{X},\mathbf{P}\otimes\nu)$. □

By adjointness the operator δ 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

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

or equivalently which decomposition (3.12) satisfies $\sum_{n=1}^{\infty} nn!\|f_n\|^2_{\mathrm{L}^2(\nu^{\otimes n})} < \infty$. The domain of δ 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 thus be extended to the respective domains of D and δ.

Proposition 3.15 (Integration by parts formula on $\mathcal{S}\times\mathcal{U}$)**.** *For any* $\mathrm{F}\in\mathbf{D}$*,* $u \in \mathrm{dom}\ \delta$*,*

$$\mathbf{E}[\mathrm{F}\,\delta(u)] = \mathbf{E}\big[\langle \mathrm{DF}, u\rangle_{\mathrm{L}^2(\mathbb{X},\tilde{\nu})}\big]. \tag{3.23}$$

3.4.2 The Ornstein-Uhlenbeck structure

Define the *Ornstein-Uhlenbeck semi-group* by its action on the chaos decomposition: for any $\mathrm{F}\in\mathrm{L}^2(\mathbf{P})$ decomposed as (3.15),

$$\mathrm{P}_\tau \mathrm{F} = \sum_{n\geqslant 0} e^{-n\tau}\mathrm{J}_n(f_n).$$

Proposition 3.16. *The Ornstein-Uhlenbeck generator (also called number operator) is defined for* $\mathrm{J}_n(f_n)\in\mathcal{H}_n$ *by*

$$\mathrm{L}(\mathrm{J}_n(f_n)) = -n\mathrm{J}_n(f_n).$$

The domain of L *fulfils the identity:* $\mathrm{F}\in\mathrm{dom}\ \mathrm{L}$ *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 $f_n \in \mathrm{L}^2(\nu^{\circ n})$, using (3.21) and then extended to $\mathrm{dom}\ \mathrm{L}$ by closability of the operator D. □

The *pseudo-inverse* L^{-1} is defined on the subspace of $\mathrm{L}^2(\mathbf{P})$ of centered random variables is given, for any F written as (3.15), by

$$\mathrm{L}^{-1}\mathrm{F} = -\sum_{n\geqslant 1}\frac{1}{n}\mathrm{J}_n(f_n). \tag{3.24}$$

3.4.3 Link with operators in L^1

The analogues of the former operators can be defined in L^1 starting from a Mecke-type formula. Let the mapping $\pi_t : \mathfrak{N}_{\mathbb{X}} \to \mathfrak{N}_{\mathbb{X}}$ be the restriction of η to $\mathcal{G}_t := \sigma\{\eta_s, s\neq t\}$, i.e.,

$$\pi_t(\eta) = \sum_{s\neq t}\sum_{k\in\mathrm{E}}\eta(s,k). \tag{3.25}$$

Lemma 3.17. *Let* η *be a marked binomial process on* $\mathbb{X}$ *with intensity measure* ν*. Then for any real-valued, non-negative,* $\mathbb{X}\times\mathfrak{N}_{\mathbb{X}}$*-measurable function* $\mathfrak{u}$*,*

$$\mathbf{E}\Bigg[\sum_{(t,k)|\eta(t,k)=1}\mathfrak{u}(\eta,(t,k))\Bigg] = \mathbf{E}\Bigg[\int_{\mathbb{X}}\mathfrak{u}\big(\pi_t(\eta)+\delta_{(t,k)},(t,k)\big)\mathrm{d}\nu(t,k)\Bigg]. \tag{3.26}$$

Proof. Formula (3.26) can be stated using similar arguments as in [16], section 2.3. □

Remark 3.18. Clearly, the formula (3.26) still holds provided the process u of representative $\mathfrak{u}$ belongs to $\mathrm{L}^1(\mathbf{P}\otimes\nu)$. Furthermore, replacing η by $\eta-\delta_{(t,k)}$ in (3.26), we can state that

$$\mathbf{E}\Bigg[\sum_{(t,k)|\eta(t,k)=1}\mathfrak{u}\big(\eta-\delta_{(t,k)},(t,k)\big)\Bigg] = \mathbf{E}\Bigg[\int_{\mathbb{X}}\mathfrak{u}\big(\pi_t(\eta),(t,k)\big)\mathrm{d}\nu(t,k)\Bigg]. \tag{3.27}$$

Gradient vs add-one cost operator

The mappings defined on $\mathfrak{N}_{\mathbb{X}} \times \mathbb{X}$ and $\mathfrak{N}_{\mathbb{X}} \times \mathbb{N}$ by

$$(\eta,(t,k)) \mapsto \pi_t(\eta)+\delta_{(t,k)} \quad \text{and} \quad (\eta,t) \mapsto \pi_t(\eta), \tag{3.28}$$

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

$$\mathrm{D}^+_{(t,k)}\mathrm{F} := \mathfrak{f}(\pi_t(\eta)+\delta_{(t,k)}) - \mathfrak{f}(\pi_t(\eta)). \tag{3.29}$$

The difference operator D^+ measures the effect of adding a point $(t,k) \in \mathbb{X}$ to η compared to the process shortened to what occurs at time t. The product formula can be deduced from (3.29) and is reminiscent to Poisson setting (see e.g. [46], Proposition 6.4.8). For $\mathrm{F},\mathrm{G} \in \mathcal{L}^0(\Omega)$ of respective representatives $\mathfrak{f}$ and $\mathfrak{g}$, such that $\mathrm{F}(\mathrm{D}^+\mathrm{G})$, $\mathrm{G}(\mathrm{D}^+\mathrm{F})$, $(\mathrm{D}^+\mathrm{F})(\mathrm{D}^+\mathrm{G}) \in \mathrm{L}^1(\mathbf{P}\otimes\nu)$,

$$\mathrm{D}^+_{(t,k)}(\mathrm{FG}) = \mathfrak{f}(\pi_t(\eta))(\mathrm{D}^+_{(t,k)}\mathrm{G}) + \mathfrak{g}(\pi_t(\eta))(\mathrm{D}^+_{(t,k)}\mathrm{F}) + (\mathrm{D}^+_{(t,k)}\mathrm{F})(\mathrm{D}^+_{(t,k)}\mathrm{G}). \tag{3.30}$$

With additional hypotheses, operators D^+ and D coincide. The definition of $\mathbf{D}$ and the chaos decomposition ensure that if $\mathrm{F} \in \mathbf{D}$, then $\mathrm{DF} \in \mathrm{L}^2(\mathbf{P})$. The following proposition provides the converse, and a more tractable expression of the gradient as a difference operator.

Proposition 3.19. *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,*

$$\mathrm{DF} = \mathrm{D}^+\mathrm{F}\ ; \quad \mathbf{P}\otimes\nu-\text{a.s.} \tag{3.31}$$

Proof. Let $(t,k) \in \mathbb{X}_T$. The mapping of $\mathrm{D}^+_{(t,k)}$ at $\mathrm{F} = \mathrm{J}_n(f_n) \in \mathcal{S}$ gives

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

where $\mathbb{X}_T^{\neg t} = \mathbb{X}_T \setminus \{(t,k),\, k \in \mathrm{E}\}$. Thus, for any $\mathrm{F} \in \mathcal{S}$, $\mathrm{D}_{(t,k)}\mathrm{F} = \mathrm{F}(\pi_t(\eta)+\delta_{(t,k)}) - \mathrm{F}(\pi_t(\eta))$. It extends to $\mathbf{D}$ by a denseness argument relying on the closability of D (Corollary 3.14). ☐

Ornstein-Uhlenbeck generator vs operator $\widetilde{\mathrm{L}}$ In this section, we study whether – under possible additional assumptions – the operators δ, L and their L^1-versions coincide. 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 $\mathfrak{u}$,

$$\tilde{\delta}(u) := \sum_{(t,k)\mid\eta(t,k)=1} \mathfrak{u}(\eta,(t,k)) - \int_{\mathbb{X}} \mathfrak{u}(\eta,(t,k))\mathrm{d}\nu(t,k) = \sum_{(t,k)\in\mathbb{X}} u_{(t,k)}\Delta\mathrm{Z}_{(t,k)}. \tag{3.32}$$

As $\pi_t(\eta) + \delta_{(t,k)} = \eta$ on $\{\eta(t,k) = 1\}$, we can additionally introduce the operator $\widetilde{\mathrm{L}}$ acting on the elements $\mathrm{F} \in \mathcal{L}^0(\Omega)$ such that $\mathrm{D}^+\mathrm{F} \in \mathrm{L}^1(\mathbf{P} \otimes \nu)$, in the following way:

$$\begin{aligned}
\widetilde{\mathrm{L}}\mathrm{F} &:= -\tilde{\delta}(\mathrm{D}^+\mathrm{F}) \\
&= -\sum_{(t,k)|\eta(t,k)=1} \big[\mathfrak{f}(\pi_t(\eta) + \delta_{(t,k)}) - \mathfrak{f}(\eta)\big] + \int_{\mathbb{X}} \big[\mathfrak{f}(\pi_t(\eta) + \delta_{(t,k)}) - \mathfrak{f}(\pi_t(\eta))\big] \mathrm{d}\nu(t,k) \\
&= -\sum_{(t,k)|\eta(t,k)=1} \big[\mathfrak{f}(\eta) - \mathfrak{f}(\eta - \delta_{(t,k)})\big] + \int_{\mathbb{X}} \big[\mathrm{D}^+_{(t,k)}\mathrm{F}\big] \,\mathrm{d}\nu(t,k) \\
&= \sum_{(t,k)|\eta(t,k)=1} \big[\mathrm{D}^-_{(t,k)}\mathrm{F}\big] + \int_{\mathbb{X}} \big[\mathrm{D}^+_{(t,k)}\mathrm{F}\big] \,\mathrm{d}\nu(t,k), \qquad (3.33)
\end{aligned}$$

where D^- can be interpreted as a *remove-one gain operator* defined for any $\mathrm{F} \in \mathcal{L}^0(\Omega)$ by

$$\mathrm{D}^-_{(t,k)}\mathrm{F} := [\mathfrak{f}(\eta) - \mathfrak{f}(\eta - \delta_{(t,k)})]\mathbf{1}_{\{\eta(t,k)=1\}}. \qquad (3.34)$$

The Mecke equation (3.26) ensures that this definition does not depend $\mathbf{P}$-almost surely on the choice of the representative. Let $\overline{\mathrm{D}}$ be the operator defined on $\mathcal{L}^0(\Omega)$ by

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

We get the following formula which is not a full L^1-integration by parts formula because of the predictability assumption.

Proposition 3.20. *For any predictable process* $u \in \mathcal{L}^0(\Omega \times \mathbb{N})$ *and* $\mathrm{F} \in \mathcal{L}^0(\Omega)$ *such that* $\mathrm{D}^+\mathrm{F}u$ *and* $\overline{\mathrm{D}}_t\mathrm{F}u$ *belong to* $\mathrm{L}^1(\mathbf{P} \otimes \nu)$*, we have*

$$\mathbf{E}\bigg[\int_{\mathbb{X}} (\mathrm{D}^+\mathrm{F})\, u_{(t,k)} \,\mathrm{d}\nu(t,k)\bigg] = \mathbf{E}[\mathrm{F}\tilde{\delta}(u)] + \mathbf{E}\bigg[\int_{\mathbb{X}} (\overline{\mathrm{D}}_t\mathrm{F})\, u_{(t,k)} \,\mathrm{d}\nu(t,k)\bigg].$$

This latter expression can be

$$\mathbf{E}\bigg[\int_{\mathbb{X}} (\widetilde{\mathrm{D}}\mathrm{F})\, u_{(t,k)} \,\mathrm{d}\nu(t,k)\bigg] = \mathbf{E}[\mathrm{F}\tilde{\delta}(u)], \qquad (3.35)$$

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. Since the intensity measure η is not diffuse, D^+ and $\widetilde{\mathrm{D}}$ are not equal $\mathbf{P} \otimes \nu$-almost surely. We end up this section with some remarks of interest for applications in Sections 5 and 6.

Conclusion: L^1 vs L^2 In the case $\mathrm{E} = \{1\}$, i.e., η is a *simple* binomial process, we have $\Delta\mathrm{Z}_t := \Delta\mathrm{Z}_{(t,1)} = \Delta\mathrm{R}_{(t,1)}$ so that $\delta = \tilde{\delta}$ and $\mathrm{LF} = \widetilde{\mathrm{L}}\mathrm{F}$. In the general case, this is no longer true. Unfortunately, even if $\mathrm{F} \in \mathrm{dom}\,\mathrm{L}$ such that $\mathrm{D}^+\mathrm{F} \in \mathrm{L}^2(\mathbf{P} \otimes \nu)$ and $\mathrm{DF} \in \mathrm{L}^1(\mathbf{P} \otimes \nu)$, we can not state $\mathrm{LF} = \widetilde{\mathrm{L}}\mathrm{F}$ $\mathbf{P}$-almost surely. Although we have $\mathrm{D}^+\mathrm{F} = \mathrm{DF}$, $\delta(\mathrm{DF}) = \sum_{(t,k)\in\mathbb{X}} (\mathrm{D}_{(t,k)}\mathrm{F})\Delta\mathrm{R}_{(t,k)}$ whereas if $\mathrm{DF} \in \mathrm{L}^1(\mathbf{P} \otimes \nu)$, $\tilde{\delta}(\mathrm{DF}) = \sum_{(t,k)\in\mathbb{X}} (\mathrm{D}_{(t,k)}\mathrm{F})\Delta\mathrm{Z}_{(t,k)}$. Nevertheless, it follows from the definition (3.3) of $\mathfrak{M}$ that

$$\mathrm{LF} = -\sum_{(t,k)\in\mathbb{X}} (\mathrm{D}_{(t,k)}\mathrm{F}) \sum_{\ell\in\mathrm{E}} \mathfrak{m}^{-1}_{k\ell}\Delta\mathrm{Z}_{(t,\ell)} = \sum_{(t,\ell)\in\mathbb{X}} \sum_{k\in\mathrm{E}} \mathfrak{m}^{-1}_{k\ell}(\mathrm{D}_{(t,k)}\mathrm{F})\Delta\mathrm{Z}_{(t,\ell)} =: \widetilde{\mathrm{L}}\widetilde{\mathrm{F}}, \qquad (3.36)$$

where $\widetilde{\mathrm{F}}$ is any square-integrable random variable such that $\mathrm{D}_{(t,\ell)}\widetilde{\mathrm{F}} = \sum_{k\in\mathrm{E}} \mathfrak{m}^{-1}_{k\ell}(\mathrm{D}^+_{(t,k)}\mathrm{F})$ for $(t,\ell) \in \mathbb{X}$. By Clark's formula (see section 4), this is uniquely defined provided $\mathbf{E}[\widetilde{\mathrm{F}}]$ is given. However, we can not combine L^1 and L^2 theories as possible in the Poisson case (see [17]) via the correspondence between the *carré du champ operator* and a L^1-operator denoted by Γ_0.

4 Some useful functional identities

4.1 Stroock's formula

The operator D^+ can be canonically iterated by letting $\mathrm{D}^{(1)} = \mathrm{D}^+$ and defining the n-th ($n \in \mathbb{N}$) *difference operator* by the recursion formula $\mathrm{D}^{(n)} = \mathrm{D}^+(\mathrm{D}^{(n-1)})$. We get explicitly for any $\mathrm{F} \in \mathcal{L}^0(\Omega)$,

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

where $\pi^{[\mathbf{t}_n]} : \mathfrak{N}_{\mathbb{X}} \to \mathfrak{N}_{\mathbb{X}}$ is the restriction of η to $\sigma\{\eta_s, s \in \mathbb{N}\setminus\{t_1,\dots,t_n\}\}$, i.e.,

$$\pi^{[\mathbf{t}_n]}(\eta) = \sum_{s\in\mathbb{N}\setminus\{t_1,\dots,t_n\}}\sum_{k\in\mathrm{E}}\eta(s,k). \tag{4.1}$$

We can then prove the discrete analogue of the *Stroock's formula*, proved by the eponymous author in the Brownian case [56]; that is the expression of the functions f_n in (3.15) in terms of the n-th difference operator $\mathrm{D}^{(n)}$. This is the exact analogue of Poisson Fock space representation.

Theorem 4.1 (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* $n \in \mathbb{N}$*, and* F *has a chaos decomposition of the form* (3.8) *with* $f_0 = \mathbf{E}[\mathrm{F}]$ *and for all* $(\mathbf{t}_n,\mathbf{k}_n) \in \mathbb{X}^n$*,*

$$f_n((\mathbf{t}_n,\mathbf{k}_n)) = \frac{1}{n!}\mathbf{E}\big[\mathrm{D}^{(n)}_{(\mathbf{t}_n,\mathbf{k}_n)}\mathrm{F}\big]. \tag{4.2}$$

Proof. This is based on the two following lemmas whose proofs are postponed to Appendix A.

Lemma 4.2. *For any* $\mathrm{F} \in \mathrm{L}^2(\mathbf{P})$*,*

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

Lemma 4.3. *For any* $\mathrm{F},\mathrm{G} \in \mathrm{L}^2(\mathbf{P})$*,*

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

The proof follows closely that of Theorem 1.3 in [32]. Let $\mathrm{F} \in \mathrm{L}^2(\mathbf{P})$ and the mapping θ_n^{F} such that:

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

Let us first state the identity for any random variable of the form $\mathrm{G} = \xi(g)$ with $g \in \mathrm{L}^2(\nu)$, well chosen to approximate F. Indeed follows from the orthogonality of the $\Delta\mathrm{R}$, that

$$\begin{aligned}\mathbf{E}\Big[\mathrm{G}\prod_{i=1}^{m}\frac{\Delta\mathrm{R}_{(s_i,\ell_i)}}{\kappa_i}\Big] = \mathbf{E}\Big[\Big(\mathbf{E}[\mathrm{G}] + \sum_{n\geqslant 1}\sum_{(\mathbf{t}_n,\mathbf{k}_n)\in\mathbb{X}^n} g_n(\mathbf{t}_n,\mathbf{k}_n)\prod_{i=1}^{n}\Delta\mathrm{R}_{(t_i,k_i)}\Big)\prod_{j=1}^{m}\frac{\Delta\mathrm{R}_{(s_j,\ell_j)}}{\kappa_j}\Big]\\ = g_m(\mathbf{s}_m,\mathbf{l}_m),\end{aligned}$$

for all $(\mathbf{s}_m,\mathbf{l}_m) \in \mathbb{X}^m$. By Lemma 4.2, the left member is also equal to $(m!)^{-1}\theta_m^{\mathrm{G}}((\mathbf{s}_m,\mathbf{l}_m))$. Now, from Lemma 4.3 together with the isometry identity (3.5), follows

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

Following the proof of Theorem 1.3 in [32] we can prove that: (i) $\mathrm{F} = \mathrm{S}$ stands $\mathbf{P}$-almost surely where $\mathrm{S} := \sum_{n\geqslant 0}(n!)^{-1}\mathrm{J}_n(\theta_n^{\mathrm{F}})$, (ii) the uniqueness of the decomposition. The proof is complete. □

4.2 Clark's formula and corollaries

The Brownian martingale representation theorem (e.g. Proposition 1.3.14 in [38]) states that a martingale adapted to the filtration of a Brownian motion is a stochastic integral whose integrand is given by Clark's formula in terms of the Malliavin gradient of the terminal value of the martingale. Here we have the analogue.

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

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

Applying the latter formula to $\mathrm{F} - \mathbf{E}\big[\mathrm{F}|\mathcal{F}_t\big]$ leads to the following result.

Corollary 4.5. *For any* $t \in \mathbb{N}$ *and* $\mathrm{F} \in \mathrm{L}^2(\mathbf{P})$*,*

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

Example 4.6. Let us introduce the normalized family $\overline{\mathcal{R}} := \{\Delta\overline{\mathrm{R}}_{(t,k)},\, (t,k) \in \mathbb{X}\}$ where $\Delta\overline{\mathrm{R}}_{(\cdot,k)} := \Delta\mathrm{R}_{(\cdot,k)}/\sqrt{\kappa_k}$ and define $\overline{\mathrm{Y}}_t = \sum_{s\in[t]}\sum_{k\in\mathrm{E}} \Delta\overline{\mathrm{R}}_{(\cdot,k)}$. Since Y_t has finite fourth moment, the application of Clark's formula (4.3) to $\overline{\mathrm{Y}}_t^2$ provides the existence of a predictable process φ such that

$$\overline{\mathrm{Y}}_t^2 = t + \sum_{s=1}^{t} \varphi_s \Delta\overline{\mathrm{Y}}_s.$$

This means that $(\overline{\mathrm{Y}}_t)_{t\in\mathbb{N}}$ satisfies a *structure equation* and then is a normal martingale. We can see again (see Section 3 and Remark 3.2) that our construction is intrinsically related to the existence of normal martingales.

We can state the analogue of the so-called Chernoff-Nash-Poincaré inequality of Gaussian analysis in [13, 35]. Our result is clearly a reminiscence of its counterpart in the Poisson space in [32, 58], for independent random variables in [15, 18, 19] or for marked point processes in [21].

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

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

Proof. The proof is similar to the one of Corollary 3.5 in [15] and will not be detailed. □

Proof of Theorem 4.4. The proof is based on the following lemma whose proof is strictly the same as Proposition 1.2.3 in [46] and will not be further detailed.

Lemma 4.8. *For any* $(t,n) \in \mathbb{N}^2$*,* $f_n \in \mathrm{L}^2(\nu^{\circ n})$*,*

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

Let $\mathrm{F} \in \mathcal{S}$. It follows from both its chaos decomposition (3.15) and the definition of the gradient operator (3.18) that for some $T \in \mathbb{N}$,

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

where we used Lemma 4.8 to obtain the second line. The operator that maps $\mathrm{F} \in \mathrm{L}^2(\mathbf{P})$ to $\big(\mathbf{E}[\mathrm{D}_{(t,k)}\mathrm{F} \,|\, \mathcal{F}_{t-1}],\, (t,k) \in \mathbb{X}\big)$ is bounded with norm equal to 1. Indeed, from (4.3) stated for cylindrical functionals together with the isometry property (3.9),

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

with equality in case $\mathrm{F} = \mathrm{J}_1(f_1)$ for some $f_1 \in \mathrm{L}^2(\nu)$. Then result can be thus extended to any random variable $\mathrm{F} \in \mathrm{L}^2(\mathbf{P})$ using a standard Cauchy argument. □

4.3 Mehler's formula

In this part, we give a pathwise representation of the pseudo-inverse L^{-1} using the operators $(\mathrm{P}_\tau)_{\tau\in\mathbb{R}_+}$. We call it *Mehler's formula* as in [31] and proceed in a similar fashion by providing an integral representation of $(\mathrm{P}_\tau)_{\tau\in\mathbb{R}_+}$ in $\mathrm{L}^1(\mathbf{P})$.

Integral representation of the semi-group For $\eta \in \mathfrak{N}_{\mathbb{X}}$, we consider the binomial process N associated to η and split it into two processes according to independent random draws of a Bernoulli random variable with mean γ. This means that any point charged by η belongs to $\eta^{(\gamma)}$ with probability γ and to $\eta^{(1-\gamma)}$ with probability $1-\gamma$. Crucially: since the measure ν is not diffuse, we need to ensure that on $\{\eta(t,k)=1\}$, the point $(t,k) \in \mathbb{X}$ can not be simultaneously charged by $\eta^{(\gamma)}$ and $\eta^{(1-\gamma)}$. For $\tau > 0$, define the probability kernel $\mathrm{K}^\tau : \{0,1\} \to [0,1]$ such that for all $(t,k) \in \mathbb{X}$, $\mathrm{K}^\tau := e^{-\tau}\delta_0 + (1-e^{-\tau})\delta_1$. Considering η as a proper process via Definition (2.1), let η_{K^τ} be the *τ-thinning* (see [33], definition 5.3) of η defined by

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

where $(\varepsilon_t^\tau)_{t\geqslant 1}$ is a sequence of variables which conditional distribution, given $\{\mathrm{N}=n\}$ (for $n \in \mathbb{N}$) **and** $\{(\mathrm{T}_t,\mathrm{V}_t),\, t \in [n]\}$, is that of independent random variables written as $\mathbf{1}_{\{\theta_t \leqslant \tau\}}$ where $(\theta_t)_{t\geqslant 1}$ is a sequence of independent exponential random variables with mean 1. We can prove that η_{K^τ} is a marked binomial process on $\mathbb{X}\times\{0,1\}$ of intensity measure $\nu\otimes\mathrm{K}^\tau$. Define

$$\eta^{\tau,0} := \eta_{\mathrm{K}^\tau}(\cdot\times\{0\}) \quad \text{and} \quad \eta^{\tau,1} := \eta_{\mathrm{K}^\tau}(\cdot\times\{1\}), \tag{4.5}$$

that are (not independent) marked binomial processes with respective intensities $e^{-\tau}\nu$ and $(1-e^{-\tau})\nu$. To see that the two processes are not independent, one can use the Laplace characterisation of binomial processes, that can be found in [33] (exercise 3.5). Nevertheless, we have $\eta^{\tau,0} + \eta^{\tau,1} = \eta$.

The formula below is very similar to the one existing in the Poisson space (see [31] or in its original formulation in [46], Lemma 6.8.1). The main difference lies in the presence here of the random variable ε. Implicitly defined in the thinning appearing in the analogue formula for Poisson processes, it is explicitly required here to guarantee that a same point can not be lighted simultaneously by $\eta^{\tau,0}$ and $\tilde{\eta}$.

Proposition 4.9. *Let $\eta \in \widehat{\mathfrak{N}}_{\mathbb{X}}$ and $\mathrm{F} \in \mathrm{L}^1(\mathbf{P})$ of representative $\mathfrak{f}$. For any $\tau \in \mathbb{R}_+$,*

$$\mathrm{P}_\tau\mathrm{F} = \mathrm{P}_\tau\mathfrak{f}(\eta^{\tau,0}+\eta^{\tau,1}) = \int \mathbf{E}\big[\mathfrak{f}(\eta^{\tau,0}+\varepsilon^\tau\tilde{\eta})\big|\eta\big]\,\Pi_\nu(\mathrm{d}\tilde{\eta})\ ;\ \mathbf{P}\text{−a.s.}, \tag{4.6}$$

where Π_ν denotes the distribution of a marked binomial process of intensity measure ν and $\tilde{\eta}$ is a point process whose law given η is such that

$$\begin{aligned}
&\mathbf{P}(\{\tilde{\eta}(t,k)=1\}\,|\,\{\eta(t,k)=0\}) = \lambda\mathbf{Q}(\{k\}) \quad \textit{and} \\
&\mathbf{P}(\{\tilde{\eta}(t,k)=0\}\,|\,\{\eta(t,k)=1\}) = 1-\lambda\mathbf{Q}(\{k\}).
\end{aligned} \tag{4.7}$$

The first equality in (4.6) ensures that for any $\mathrm{F} \in \mathrm{L}^1(\mathbf{P})$, $\tau \in \mathbb{R}_+$,

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

while Jensen's inequality together with (4.6) imply the contractivity property of $(\mathrm{P}_\tau)_{\tau \in \mathbb{R}_+}$:

$$\mathbf{E}[|\mathrm{P}_\tau \mathrm{F}|^p] \leqslant \mathbf{E}[|\mathrm{F}|^p] \; ; \; p \in \mathbb{N}. \tag{4.8}$$

Proof of Proposition 4.9. It is enough to prove it for $\mathrm{F} = \xi(h) = \mathfrak{f}(\eta)$ where, as η is finite, there exists $T \in \mathbb{N}$ such that

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

with $g \in \mathrm{L}^2(\mathbb{X}_T)$ 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-chaos decomposition (3.15),

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

On the other hand, by definition of $\eta^{\tau,0}$ (4.5) and $\tilde{\eta}$, whose law given η is provided by (4.7),

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

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

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

Corollary 4.10. *For any* $\mathrm{F} \in \mathrm{L}^2(\mathbf{P})$*, and* $\tau \in \mathbb{R}_+$*,*

$$\mathrm{D}\mathrm{P}_\tau \mathrm{F} = e^{-\tau} \mathrm{P}_\tau \mathrm{D}\mathrm{F} \; ; \; \mathbf{P} - \text{a.s.} \tag{4.9}$$

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

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

where $g \in \mathrm{L}^2(\nu)$ is such that $\mathrm{J}_1(h) = \mathrm{J}_1(g\, ; \mathcal{Z})$. Then, from Mehler's formula (4.6),

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

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

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

On the other hand, follows from

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

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

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

Hence the result. □

Mehler's formula appears as a useful corollary of the commutation property.

Theorem 4.11. *For any* $\mathrm{F}\in\mathrm{L}^2(\mathbf{P})$ *such that* $\mathbf{E}[\mathrm{F}]=0$*,*

$$\mathrm{L}^{-1}\mathrm{F} = -\int_0^\infty \mathrm{P}_\tau\mathrm{F}\,\mathrm{d}\tau\ ;\ \mathbf{P}\otimes\nu - \text{a.e.} \tag{4.10}$$

Moreover,

$$-\,\mathrm{D}\mathrm{L}^{-1}\mathrm{F} = \int_0^\infty e^{-\tau}\mathrm{P}_\tau\mathrm{D}\mathrm{F}\,\mathrm{d}\tau\ ;\ \mathbf{P}\otimes\nu - \text{a.e.} \tag{4.11}$$

The combination of Theorem 4.11 with the contraction property of $(\mathrm{P}_\tau)_{t\in\mathbb{R}_+}$ enables to bound $\mathrm{D}\mathrm{L}^{-1}\mathrm{F}$ with respect to the norm of $\mathrm{D}\mathrm{F}$:

$$\|\mathrm{D}\mathrm{L}^{-1}\mathrm{F}\|_{\mathrm{L}^2(\mathbf{P}\otimes\nu)} \leqslant \|\mathrm{D}\mathrm{F}\|_{\mathrm{L}^2(\mathbf{P}\otimes\nu)}.$$

Proof of Theorem 4.11. Let $\mathrm{F}\in\mathrm{L}^2(\mathbf{P})$, $\mathbf{E}[\mathrm{F}]=0$. By (4.8), $\mathrm{P}_\tau\mathrm{F}\in\mathrm{L}^2(\mathbf{P})$ for any $\tau\in\mathbb{N}$.

$$\mathrm{L}^{-1}\left(\sum_{n=1}^m \mathrm{J}_n(f_n)\right) = -\sum_{n=1}^m \frac{1}{n}\mathrm{J}_n(f_n) = -\int_0^\infty \sum_{n=1}^m e^{-n\tau}\mathrm{J}_n(f_n)\,\mathrm{d}\tau, \tag{4.12}$$

for any $m\in\mathbb{N}$. Moreover, the random variable 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

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

We have that

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

and then the first point is proved by letting m go to infinity in (4.12). As for (4.11), the commutation (4.9) and contractivity (4.8) properties satisfied by $(\mathrm{P}_\tau)_{\tau\in\mathbb{R}_+}$ ensure that

$$\mathbf{E}\Big[\int_0^\infty |\mathrm{D}_{(s,\ell)}\mathrm{P}_\tau\mathrm{F}|\,\mathrm{d}\tau\Big] = \mathbf{E}\Big[\int_0^\infty e^{-\tau}|\mathrm{P}_\tau\mathrm{D}_{(s,\ell)}\mathrm{F}|\,\mathrm{d}\tau\Big] \leqslant (\sqrt{2})^{-1}\mathbf{E}\big[|\mathrm{D}_{(s,\ell)}\mathrm{F}|\big]$$

is finite for ν-a.e. $(s,\ell)\in\mathbb{X}$. The result follows by applying the operator D to each side of equality (4.10) and using of the commutation property (4.9). □

5 Application 1: Poisson approximations via Stein's method

The Chen-Stein method (see e.g. [8, 20]) was derived from Stein's (see e.g. [5]) to assess probability distances between the law of an integer-valued random variable and a Poisson distribution. The first related results dealt with the Poisson approximation of sums of possibly dependent random variables (see [1]) and were based on the study of *neighbourhoods of dependence*. In this section, we revisit this latter by combining it with the Malliavin calculus for MBP. In particular, we state a Stein-Malliavin type criterion (see e.g. [36, 37] for the historical result for Normal approximation) for the Poisson (resp.compound Poisson) approximation by simple binomial (resp. marked binomial) functionals with respect to the total variation distance. This is defined for two $\mathbb{Z}_+$-random variables X and Y (the case of interest here), by

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

where for any random variable X, we denote by $\mathbf{P}_\mathrm{X}$ its distribution, i.e., the pushforward measure of $\mathbf{P}$ by X. First, we can state a result for Poisson approximation, in the same vein as [43] on the Poisson space. Let $\mathcal{P}(\lambda_0)$ the Poisson law with mean λ_0. For a function $\varphi : \mathbb{Z}_+ \to \mathbb{R}$, consider $\nabla\varphi$ 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 (see e.g. [43], proof of Theorem 3.3) inequality: for all $a, k \in \mathbb{Z}_+$,

$$\big|\varphi(k) - \varphi(a) - \nabla\varphi(a)(k-a)\big| \leqslant \frac{\|\nabla^2\varphi\|_\infty}{2}|(k-a)(k-a-1)|. \tag{5.1}$$

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

$$\mathbf{P}(\mathcal{P}(\lambda_0)\in\mathrm{A}) - \mathbf{1}_\mathrm{A}(k) = k\varphi_\mathrm{A}(k) - \lambda_0\varphi_\mathrm{A}(k+1)\ ;\ k\in\mathbb{Z}_+, \tag{5.2}$$

satisfying the condition $\nabla^2\varphi_\mathrm{A}(0) = 0$. Let us denote $\|\varphi\|_\infty = \max_{\mathrm{A}\subset\mathbb{Z}_+}\|\varphi_\mathrm{A}\|_\infty$, $\|\nabla\varphi\|_\infty = \max_{\mathrm{A}\subset\mathbb{Z}_+}\|\nabla\varphi_\mathrm{A}\|_\infty$ and $\|\nabla^2\varphi\|_\infty = \max_{\mathrm{A}\subset\mathbb{Z}_+}\|\nabla^2\varphi_\mathrm{A}\|_\infty$. The function class $\mathcal{K} = \{\varphi_\mathrm{A},\ \mathrm{A}\subset\mathbb{Z}_+\}$ fulfils the estimates (see [43]):

$$\|\varphi\|_\infty \leqslant \min\Big(1, \sqrt{\frac{2}{e\lambda_0}}\Big),\ \|\nabla\varphi\|_\infty \leqslant \frac{1-e^{-\lambda_0}}{\lambda_0}, \tag{5.3}$$

as well as

$$\|\nabla^2\varphi\|_\infty \leqslant 2\|\nabla\varphi\|_\infty \leqslant \frac{2(1-e^{-\lambda_0})}{\lambda_0}. \tag{5.4}$$

Note that this last bound is connected to Remark 4.4 in Torrisi ([57], page 2227) where the need to replace the bound $(1-e^{-\lambda_0})/\lambda_0^2$ in Theorem 3.1 of [43] by $(1-e^{-\lambda_0})/\lambda_0$ has been pointed out.

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

$$\begin{aligned}\mathrm{d_{TV}}\big(\mathbf{P}_\mathrm{F}, \mathcal{P}(\lambda_0)\big) \leqslant \frac{1-e^{-\lambda_0}}{\lambda_0}&\sqrt{\mathbf{E}\Big[\Big|\lambda_0 - \int_\mathbb{N}(\widetilde{\mathrm{D}}\mathrm{F})\mathrm{D}\mathrm{L}^{-1}(\mathrm{F}-\mathbf{E}[\mathrm{F}])\nu(\mathrm{d}t)\Big|^2\Big]} \\ &+ \frac{1-e^{-\lambda_0}}{\lambda_0}\mathbf{E}\Big[\int_\mathbb{N}\big|(\widetilde{\mathrm{D}}_t\mathrm{F})(\widetilde{\mathrm{D}}_t\mathrm{F}-1)\big|\big|\mathrm{D}_t\mathrm{L}^{-1}(\mathrm{F}-\mathbf{E}[\mathrm{F}])\big|\,\nu(\mathrm{d}t)\Big].\end{aligned} \tag{5.5}$$

Proof. Let $\mathrm{A}\subset\mathbb{Z}_+$ and φ_A be the solution of the Stein equation (5.2). First, letting $u = \mathrm{D}^+\mathrm{L}^{-1}(\mathrm{F}-\mathbf{E}[\mathrm{F}])$ and $\mathrm{G} = \varphi_\mathrm{A}(\mathrm{F})$, we have

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

This implies that $(\mathrm{D}_t^+\mathrm{G})u \in \mathrm{L}^1(\mathbf{P}\otimes\nu)$ since $\mathrm{D}^+\mathrm{L}^{-1}(\mathrm{F}-\mathbf{E}[\mathrm{F}]) \in \mathrm{L}^1(\mathbf{P}\otimes\nu)$ as a consequence of (4.11) together with the contraction property (4.8). Similarly we can prove that $(\overline{\mathrm{D}}_t\mathrm{G})u \in \mathrm{L}^1(\mathbf{P}\otimes\nu)$ so that the integration by parts formula (3.35) holds. Moreover, since $\mathrm{F}\in\mathbf{D}$, then $\mathrm{DF} = \mathrm{D}^+\mathrm{F}$ $\mathbf{P}$-almost surely and, as $\mathrm{E}=\{1\}$, it follows from last part of Section 3 that $\widetilde{\delta}=\delta$ and $\widetilde{\mathrm{L}}\mathrm{F} = \mathrm{LF}$. With our notation (see Section 2), we can write here that $\pi_t(\eta)+\delta_t = \eta+\delta_t$ $\mathbf{P}$-almost surely so that $\widetilde{\mathrm{D}}_t\mathrm{F} = \mathfrak{f}(\eta+\delta_t)-\mathfrak{f}(\eta)$ $\mathbf{P}$-almost surely. By definition of the operators L and L^{-1} and using (3.35), we get

$$\begin{aligned}
\mathbf{E}\Big[\mathrm{F}\varphi_{\mathrm{A}}(\mathrm{F})-\lambda_0\varphi_{\mathrm{A}}(\mathrm{F}+1)\Big] &= \mathbf{E}\Big[(\mathrm{F}-\mathbf{E}[\mathrm{F}])\varphi_{\mathrm{A}}(\mathrm{F})-\lambda_0\nabla\varphi_{\mathrm{A}}(\mathrm{F})\Big]\\
&= \mathbf{E}\Big[(\widetilde{\mathrm{L}}\widetilde{\mathrm{L}}^{-1}(\mathrm{F}-\mathbf{E}[\mathrm{F}]))\varphi_{\mathrm{A}}(\mathrm{F})\Big]-\mathbf{E}\Big[\lambda_0\nabla\varphi_{\mathrm{A}}(\mathrm{F})\Big]\\
&= -\mathbf{E}\Big[\int_{\mathbb{N}}\widetilde{\mathrm{D}}_t(\varphi_{\mathrm{A}}(\mathrm{F}))\,\mathrm{D}_t\mathrm{L}^{-1}(\mathrm{F}-\mathbf{E}[\mathrm{F}])\,\nu(\mathrm{d}t)\Big]-\mathbf{E}\big[\lambda_0\nabla\varphi_{\mathrm{A}}(\mathrm{F})\big]\\
&= -\mathbf{E}\Big[\nabla\varphi_{\mathrm{A}}(\mathrm{F})\int_{\mathbb{N}}(\widetilde{\mathrm{D}}_t\mathrm{F})\,\mathrm{D}_t\mathrm{L}^{-1}(\mathrm{F}-\mathbf{E}[\mathrm{F}])\nu(\mathrm{d}t)+\mathrm{rem}_{\mathrm{A}}\Big]\\
&\qquad -\mathbf{E}\big[\lambda_0\nabla\varphi_{\mathrm{A}}(\mathrm{F})\big],
\end{aligned}$$

where the last line comes from

$$\begin{aligned}
\int_{\mathbb{N}}\widetilde{\mathrm{D}}_t(\varphi_{\mathrm{A}}(\mathrm{F}))\,\mathrm{D}_t\mathrm{L}^{-1}(\mathrm{F}-\mathbf{E}[\mathrm{F}])\,\nu(\mathrm{d}t) &= \nabla\varphi_{\mathrm{A}}(\mathrm{F})\int_{\mathbb{N}}\big(\widetilde{\mathrm{D}}_t\mathrm{F}\big)\big(\mathrm{D}_t\mathrm{L}^{-1}(\mathrm{F}-\mathbf{E}[\mathrm{F}])\big)\,\nu(\mathrm{d}t)\\
&\quad+\int_{\mathbb{N}}\mathfrak{R}_{\mathrm{A},t}\mathrm{D}_t\mathrm{L}^{-1}(\mathrm{F}-\mathbf{E}[\mathrm{F}])\,\nu(\mathrm{d}t),
\end{aligned}$$

and $\mathfrak{R}_{\mathrm{A},t}$ is a residual random function such that $\mathfrak{R}_{\mathrm{A},t}\leqslant\|\nabla^2\varphi_{\mathrm{A}}\|_\infty\big|(\widetilde{\mathrm{D}}_t\mathrm{F})(\widetilde{\mathrm{D}}_t\mathrm{F}-1)\big|/2$. By using inequality (5.1), with $k=\mathfrak{f}(\eta)+\widetilde{\mathrm{D}}_t\mathrm{F}$ and $a=\mathfrak{f}(\eta)$, we get

$$\begin{aligned}
\mathbf{E}\big[|\mathrm{rem}_{\mathrm{A}}|\big] &= \mathbf{E}\Big[\Big|\int_{\mathbb{N}}\mathfrak{R}_{\mathrm{A},t}(\mathrm{D}_t\mathrm{L}^{-1}(\mathrm{F}-\mathbf{E}[\mathrm{F}]))\,\nu(\mathrm{d}t)\Big|\Big]\\
&\leqslant \frac{\|\nabla^2\varphi_{\mathrm{A}}\|_\infty}{2}\mathbf{E}\Big[\int_{\mathbb{N}}|(\widetilde{\mathrm{D}}_t\mathrm{F})(\widetilde{\mathrm{D}}_t\mathrm{F}-1)|\big|\mathrm{D}_t\mathrm{L}^{-1}(\mathrm{F}-\mathbf{E}[\mathrm{F}])\big|\,\nu(\mathrm{d}t)\Big].
\end{aligned}$$

Then, from Cauchy-Schwarz inequality,

$$\begin{aligned}
\big|\mathbf{E}\big[\mathrm{F}\varphi_{\mathrm{A}}(\mathrm{F})-\lambda_0\varphi_{\mathrm{A}}(\mathrm{F}+1)\big]\big| &\leqslant \|\nabla\varphi_{\mathrm{A}}\|_\infty\sqrt{\mathbf{E}\Big[\Big|\lambda_0-\int_{\mathbb{N}}(\widetilde{\mathrm{D}}_t\mathrm{F})\,\mathrm{D}_t\mathrm{L}^{-1}(\mathrm{F}-\mathbf{E}[\mathrm{F}])\nu(\mathrm{d}t)\Big|^2\Big]}\\
&\quad+\frac{\|\nabla^2\varphi_{\mathrm{A}}\|_\infty}{2}\mathbf{E}\Big[\int_{\mathbb{N}}\big|(\widetilde{\mathrm{D}}_t\mathrm{F})(\widetilde{\mathrm{D}}_t\mathrm{F}-1)\big|\big|\mathrm{D}_t\mathrm{L}^{-1}(\mathrm{F}-\mathbf{E}[\mathrm{F}])\big|\,\nu(\mathrm{d}t)\Big].
\end{aligned}$$

The result is then obtained by taking the supremum over the set $\{\mathrm{A}\subset\mathbb{Z}_+\}$ and using the uniform bounds (5.3) on $\nabla_\cdot\varphi$ and $\nabla^2_\cdot\varphi$. ☐

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 with parameters $(\lambda_0,\mathbf{V})$, i.e., that can be written as the distribution of the variable

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

where N^{P} is a Poisson random variable with mean λ_0 and $\{\mathrm{V}_i,\, i\in\mathbb{N}\}$ is a family of independent non-negative random variables with distribution $\mathbf{V}$. For any $\mathrm{A}\subset\mathbb{Z}_+$, we denote by ψ_{A} the unique solution of the Chen-Stein equation (see [6], Theorem 1)

$$\mathbf{1}_{\mathrm{A}}(\ell)-\mathbf{P}(\mathcal{PC}(\lambda_0,\mathbf{V})) = \ell\psi_{\mathrm{A}}(\ell)-\int_{\mathbb{X}}k\psi_{\mathrm{A}}(\ell+k)\,\mathrm{d}\nu(t,k)\ ;\ \ell\in\mathbb{Z}_+. \tag{5.6}$$

The function class $\mathcal{K}' = \{\varphi_{\mathrm{A}},\ \mathrm{A} \subset \mathbb{Z}_+\}$ satisfies (see [20], Theorem 3.5):

$$\mathfrak{d}_{\mathcal{PC}} := \max_{\mathrm{A}\subset\mathbb{Z}_+} \|\psi_{\mathrm{A}}\|_\infty \vee \max_{\mathrm{A}\subset\mathbb{Z}_+} \|\nabla\psi_{\mathrm{A}}\|_\infty \leqslant \min\left(1, \frac{1}{\lambda_0 \mathbf{V}(\{1\})}\right) e^{\lambda_0}. \tag{5.7}$$

Proposition 5.2. *Let $\lambda_0 \in \mathbb{R}_+^*$, $\mathbf{V}$ a probability distribution on $\mathbb{N}$ and V_1 a random variable with law $\mathbf{V}$. For any $\mathbb{Z}_+$-valued random variable $\mathrm{F} \in \mathbf{D}$ such that $\mathbf{E}[\mathrm{F}] = \lambda_0 \mathbf{E}[\mathrm{V}_1]$,*

$$\begin{aligned} &\mathrm{d_{TV}}\big(\mathbf{P}_{\mathrm{F}}, \mathcal{PC}(\lambda_0, \mathbf{V})\big) \\ &\qquad \leqslant \Big| \int_{\mathbb{X}} \big[\mathrm{D}^+\widetilde{\mathrm{L}}^{-1}(\mathfrak{f}(\eta) - \mathbf{E}[\mathfrak{f}(\eta)])\psi_{\mathrm{A}}(\mathfrak{f}(\pi_t(\eta) + \delta_{(t,k)})) - k\psi_{\mathrm{A}}(\mathfrak{f}(\eta) + k)\big] \mathrm{d}\nu(t,k) \Big| \\ &\qquad\quad + \mathfrak{d}_{\mathcal{PC}} \Big| \int_{\mathbb{X}} \big[\mathrm{D}^+\widetilde{\mathrm{L}}^{-1}(\mathfrak{f}(\eta) - \mathbf{E}[\mathfrak{f}(\eta)]) - k\big]\, \mathrm{d}\nu(t,k) \Big|. \end{aligned} \tag{5.8}$$

This result is only interesting when the variable F is a marked binomial functional in the first chaos, i.e., $\mathrm{F} = \mathrm{J}_1(f)$ ($f \in \mathrm{L}^2(\nu)$) and is not relevant for functionals in higher chaoses. In this latter case, we can provide a bound by means of a Taylor expansion and in terms of the iterated operator ∇^2. We choose not to present it since it turns out to be sub-optimal in the case $\mathrm{F} \in \mathcal{H}_1$ that coincides with the frame of the application in Section 5.2.

Proof of Proposition 5.2. Let a $\mathbb{Z}_+$-valued random variable $\mathrm{F} \in \mathbf{D}$ such that $\mathbf{E}[\mathrm{F}] = \lambda_0 \mathbf{E}[\mathrm{V}_1]$. Via Stein's method for Poisson compound approximation, we are led to control

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

Let $u = \mathrm{D}^+\mathrm{L}^{-1}(\mathrm{F} - \mathbf{E}[\mathrm{F}])$ and $\mathrm{G} = \psi_{\mathrm{A}}(\mathrm{F})$. We have

$$|(\mathrm{D}_t^+\mathrm{G})u| \leqslant \|\varphi\|_\infty |\mathrm{D}^+\mathrm{L}^{-1}(\mathrm{F} - \mathbf{E}[\mathrm{F}])|.$$

Then, $(\mathrm{D}_t^+\mathrm{G})u \in \mathrm{L}^1(\mathbf{P}\otimes\nu)$ since $\mathrm{D}^+\mathrm{L}^{-1}(\mathrm{F} - \mathbf{E}[\mathrm{F}]) \in \mathrm{L}^1(\mathbf{P}\otimes\nu)$ from (4.11) together with the contraction property (4.8). Then, by definition of $\widetilde{\mathrm{L}}$ and $\widetilde{\mathrm{L}}^{-1}$ and using (3.35),

$$\mathbf{E}\big[(\mathrm{F} - \mathbf{E}[\mathrm{F}])\psi_{\mathrm{A}}(\mathrm{F})\big] = \mathbf{E}\big[(\widetilde{\mathrm{L}}\widetilde{\mathrm{L}}^{-1})(\mathrm{F} - \mathbf{E}[\mathrm{F}])\psi_{\mathrm{A}}(\mathrm{F})\big] = \mathbf{E}\Big[\int_{\mathbb{X}} \mathrm{D}\widetilde{\mathrm{L}}^{-1}(\mathrm{F} - \mathbf{E}[\mathrm{F}])\widetilde{\mathrm{D}}\psi_{\mathrm{A}}(\mathrm{F})\mathrm{d}\nu(t,k)\Big].$$

On the other hand,

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

We conclude by taking the expectation and then the supremum over $\{\psi_{\mathrm{A}},\ \mathrm{A} \subset \mathbb{Z}_+\}$. □

5.1 Head run problems

Consider a large number of independent coin tosses with success (falling head) probability $p_n \in (0,1)$. Whatever the value of p_n, there will be sequences, called *head runs*, where the coin will fall on head each time. We aim at computing the probability that U_n, the length of the longest head run beginning in the first n tosses, is less than a test length $m_n \in \mathbb{N}$. The crucial fact is that head runs occur in *clumps*; indeed, if there is a run with length m_n at position i, the probability that another run with the same length starts at position $i+1$ is p_n. Considering $m_n = 5$ and the series

$$0001100101010\underbrace{11111}_{m_n=5}010000\underbrace{11111}_{m_n=5}\underbrace{11111}_{m_n=5}1001$$

$$0001100101010\underbrace{11111}_{m_n=5}0100001\underbrace{11111}_{m_n=5}\underbrace{11111}_{m_n=5}001$$

we see that there are runs with length m_n at positions $j = 14, 26, 27, 28, 29, 30, 31, 32$. We need to "declump" the sequences in order to count only the first occurrence. Let $(\mathrm{C}_i^n)_{i\in\mathbb{N}}$ be a sequence of i.i.d. Bernoulli random variables with mean p_n, $m_n \in \mathbb{N}$ be the "test" value and consider the random variable

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

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

$$\mathrm{U}_n = \prod_{i=1}^{m_n} \Delta\mathrm{N}_i^n + \sum_{i=1}^{n-1} (1-\Delta\mathrm{N}_i^n)\Delta\mathrm{N}_{i+1}^n \Delta\mathrm{N}_{i+2}^n \cdots \Delta\mathrm{N}_{i+m_n}^n =: \mathrm{U}_0^n + \sum_{i=1}^{n-1} \mathrm{U}_i^n, \tag{5.9}$$

i.e., as a simple binomial functional with mean p_n. As $\mathrm{U}_n \in \mathbf{D}$, $\mathrm{D}\mathrm{U}_n = \mathrm{D}^+\mathrm{U}_n$ $\mathbf{P}\otimes\nu$-almost surely, and

$$\mathrm{D}_t\mathrm{U}_n = \mathbf{1}_{[m_n]}(t) \prod_{i=1, i\neq t}^{m_n} \Delta\mathrm{N}_i^n + \sum_{i=1}^{n-1} \Big(\mathbf{1}_{[i+1,i+m_n]}(t)(1-\Delta\mathrm{N}_i^n) \prod_{\ell=1, i+\ell\neq t}^{m_n} \Delta\mathrm{N}_{i+\ell}^n + \mathbf{1}_{\{i\}}(t) \prod_{\ell=1}^{m_n} \Delta\mathrm{N}_{i+\ell}^n \Big) \tag{5.10}$$

is not null for $t \in \mathcal{J}_n := [n+m_n-1]$. We can quantify the Poisson approximation of U_n.

Theorem 5.3. *Let $\lambda_n = p_n^m((n-1)(1-p_n)+1)$. Then, there exists $\mathfrak{c} > 0$ such that*

$$\begin{aligned} \mathrm{d_{TV}}(\mathbf{P}_{\mathrm{U}_n}, \mathcal{P}(\lambda_n)) \leqslant \frac{p_n^{m_n}}{\lambda_n} \Big[& p_n^{m_n}[2(m_n-1)q_n^2 + 2m_n q_n + 1] + \mathfrak{c} m_n^2 (1-p_n)^2 p_n^{m_n-1} \\ & + 2(n-m_n+1)(1-p_n)p_n^{m_n} + \mathrm{o}(p_n^{m_n}) \Big]. \end{aligned} \tag{5.11}$$

Remark 5.4. This result gives an insight into the distribution of T_n, the length of the longest head run. As explained in [2], the distribution of T_n may be approximated as

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

The definition of a test length requires that the sequence $(\lambda_n)_{n\geqslant 1}$ is bounded away from 0 and ∞. In other words, this means there exists a deterministic constant c such that

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

Under this assumption, $(1/\lambda_n)_{n\geqslant 1}$ is bounded and $p_n^{m_n}$ is asymptotically equivalent to $1/n$, so that the Poisson approximation in Theorem 5.3 is of order $1/n$. We do slightly better than in [2] where the convergence rate is of order m_n/n and is obtained by studying neighbourhood of dependence to deal with the local dependence structure of U_n.

Proof of Theorem 5.3. Since U is a functional of a *simple binomial process*, we have $\mathrm{D}_t\mathrm{U}_n = \mathrm{D}_t^+\mathrm{U}_n$, $\widetilde{\mathrm{D}}_t\mathrm{U}_n = \mathbf{1}_{\{\eta(t)=0\}}\mathrm{D}^+\mathrm{U}_n$, and $\widetilde{\mathrm{L}}\mathrm{U}_n = \mathrm{L}\mathrm{U}_n$. Note that $\lambda_n^2 = p_n^{2m} + 2(n-1)q_np_n^{2m} + (n-1)^2q_n^2p_n^{2m}$. We want to bound A_1 and A_2 the two terms in the right-hand side of (5.5), i.e.,

$$A_1 := \sqrt{\mathbf{E}\Big[\Big|\lambda_n - \int_{\mathcal{J}_n}(\widetilde{\mathrm{D}}_t\mathrm{U}_n)\mathrm{D}_t\mathrm{L}^{-1}(\mathrm{U}_n - \mathbf{E}[\mathrm{U}_n])\nu(\mathrm{d}t)\Big|^2\Big]},$$

and

$$A_2 := \mathbf{E}\Big[\int_{\mathcal{J}_n}\big|(\widetilde{\mathrm{D}}_t\mathrm{U}_n)(\widetilde{\mathrm{D}}_t\mathrm{U}_n - 1)\big|\big|\mathrm{D}_t\mathrm{L}^{-1}(\mathrm{U}_n - \mathbf{E}[\mathrm{U}_n])\big|\nu(\mathrm{d}t)\Big].$$

Our demonstration is based on the following lemma whose proof is postponed to Appendix B.

Lemma 5.5.

$$\sqrt{\mathbf{E}\Big[\Big|\mathrm{var}[\mathrm{F}] - \int_{\mathcal{J}_n}\widetilde{\mathrm{D}}_t\mathrm{F}(-\mathrm{D}_t\mathrm{L}^{-1}\mathrm{F})\mathrm{d}\nu(t)\Big|^2\Big]} \leqslant \mathfrak{c}m_n^2(1-p_n)^2p_n^{2m_n-1} + \mathrm{o}(p_n^{2m_n}).$$

Bound on A_1

$$A_1 \leqslant \big|\lambda_n - \mathrm{var}[\mathrm{U}_n]\big| + \sqrt{\mathbf{E}\Big[\Big|\mathrm{var}[\mathrm{U}_n] - \int_{\mathcal{J}_n}(\widetilde{\mathrm{D}}_t\mathrm{U}_n)\mathrm{D}_t\mathrm{L}^{-1}(\mathrm{U}_n - \mathbf{E}[\mathrm{U}_n])\nu(\mathrm{d}t)\Big|^2\Big]} =: A_{11} + A_{12},$$

where, using that for $j > i$, $\mathbf{E}[\mathrm{U}_i^n\mathrm{U}_j^n] = 0$ if $j \in [\![i+1, i+m_n]\!]$, we have

$$\begin{aligned}
A_{11} &= \Big|\lambda_n - \Big(\lambda_n + 2\sum_{i=m_n+1}^{n-1}\mathbf{E}[\mathrm{U}_0^n\mathrm{U}_i^n] + 2\sum_{i=1}^{n-1}\sum_{j=i+m_n+1}^{n-1}\mathbf{E}[\mathrm{U}_i^n\mathrm{U}_j^n] - \lambda_n^2\Big)\Big| \\
&= 2(n-m_n-1)q_np_n^{2m_n} + (n-1)(n-2m_n-2)q_n^2p_n^{2m_n} - \lambda_n^2 \\
&= 2m_nq_np_n^{2m_n} + (2m_n-1)q_n^2p_n^{2m_n} + p_n^{2m_n}.
\end{aligned}$$

This together with Lemma 5.5 leads to the existence of a constant $\mathfrak{c}$ such that

$$A_1 \leqslant p_n^{2m_n}[2(m_n-1)q_n^2 + 2m_nq_n + 1] + \mathfrak{c}m_n^2p_n^{2m_n-1} + \mathrm{o}(p_n^{2m_n}). \tag{5.12}$$

Bound on A_2 Using Corollary 4.11,

$$\begin{aligned}
A_2 &\leqslant \mathbf{E}\Big[\int_{\mathcal{J}_n}\Big(\mathbf{1}_{\{\eta(t)=0\}}\big|(\mathrm{D}_t\mathrm{U}_n)^2 - (\mathrm{D}_t\mathrm{U}_n)\big|\Big|\int_0^\infty \mathrm{D}_t\mathrm{P}_\tau\mathrm{U}_n\,\mathrm{d}\tau\Big|\Big)\nu(\mathrm{d}t)\Big] \\
&= (1-p_n)\mathbf{E}\Big[\int_{\mathcal{J}_n}\Big(\big|(\mathrm{D}_t\mathrm{U}_n)^2 - (\mathrm{D}_t\mathrm{U}_n)\big|\Big|\int_0^\infty e^{-\tau}\mathrm{P}_\tau\mathrm{D}_t\mathrm{U}_n\,\mathrm{d}\tau\Big|\Big)\nu(\mathrm{d}t)\Big] \\
&\leqslant (1-p_n)\mathbf{E}\Big[\int_{\mathcal{J}_n}\Big(\big|(\mathrm{D}_t\mathrm{U}_n)^2 - (\mathrm{D}_t\mathrm{U}_n)\big|\Big)\nu(\mathrm{d}t)\Big],
\end{aligned}$$

by conditioning w.r.t. $\mathcal{F}_{t-1}$ and using that $|\mathrm{D}_t\mathrm{U}_n| \leqslant 1$. Moreover,

$$(\mathrm{D}_t\mathrm{U}_n)^2 = \Big(\prod_{i=1,i\neq t}^{m_n}\Delta\mathrm{N}_i + \sum_{i=\max(t-m_n,1)}^{t-1}(1-\Delta\mathrm{N}_i)\prod_{\ell=1,i+\ell\neq t}^{m_n}\Delta\mathrm{N}_{i+\ell} - \prod_{\ell=1}^{m_n}\Delta\mathrm{N}_{t+\ell}\Big)^2 =: (\mathrm{A}+\mathrm{B}-\mathrm{C})^2,$$

Case $t > m_n$ Note that $\mathrm{D}_t\big(\prod_{i=1}^{m_n}\Delta\mathrm{N}_i\big) = 0$ so that $\mathrm{D}_t\mathrm{U}_n = \mathrm{B}-\mathrm{C}$. Assume there exists $i_0 \in \{t-m_n, t-1\}$ such that $\mathrm{B}_{i_0} = 1$ where $\mathrm{B}_i := (1-\Delta\mathrm{N}_i)\prod_{\ell=1,i+\ell\neq t}^{m_n}\Delta\mathrm{N}_{i+\ell}$. Then, $\Delta\mathrm{N}_{i_0} = 0$ implies $\mathrm{B}_i = 0$ for any $i \in [\![\max(1, i_0-m_n), \min(t-1, i_0-1)]\!] = [\![\max(1, i_0-m_n), i_0-1]\!]$ whereas $\Delta\mathrm{N}_{i_0+\ell} = 1$ leads to $\mathrm{B}_i = 0$ for any $i \in [\![\max(t-m_n, i_0+1), \min(t-1, i_0+m_n)]\!] = [\![i_0+1, t-1]\!]$; this entails $\mathrm{B} = \sum_{i=t-m_n}^{t-1}\mathrm{B}_i = 0 + \mathrm{B}_{i_0} = 1$. Thus $\mathrm{B} \in \{0,1\}$ and $\mathbf{P}(\{\mathrm{B} \leqslant 1\}) = 1$. Besides, since $\mathrm{B}, \mathrm{C} \in \{0,1\}$,

$$(\mathrm{D}_t\mathrm{U}_n)^2 = \mathrm{B}^2 - 2\mathrm{BC} + \mathrm{C}^2 = (\mathrm{B}-\mathrm{C}) - 2\mathrm{BC} + 2\mathrm{C} = (\mathrm{D}_t\mathrm{U}_n) + 2\mathrm{C}(1-\mathrm{B}).$$

This proves that $\mathrm{D}_t\mathrm{U}_n \leqslant (\mathrm{D}_t\mathrm{U}_n)^2$ $\mathbf{P}$-a.s. On the event $\{\mathrm{C}=1\}$, $\prod_{\ell=1}^{m_n}\Delta\mathrm{N}_{t+\ell} = 1$ so that B_{t-1} is equal to 0 if and only if $\Delta\mathrm{N}_{t-1} = 1$. This entails $\prod_{\ell=1,\ell\neq 2}^{m_n}\Delta\mathrm{N}_{t-2+\ell}$ is equal to 1 and then 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 $i \in \{t-m_n, \ldots, t-1\}$ (with probability $p_n^{m_n}$). Then

$$\mathbf{E}[\mathrm{C}(1-\mathrm{B})] = \mathbf{E}[\mathbf{1}_{\{\mathrm{B}=0\}}\big|\{\mathrm{C}=1\}]\mathbf{P}(\{\mathrm{C}=1\}) = p_n^{2m_n}.$$

so that we get for $t > m_n$,

$$\mathbf{E}[(\mathrm{D}_t\mathrm{U}_n)^2 - \mathrm{D}_t\mathrm{U}_n] = p_n^{2m_n},$$

Moreover, for $t \leqslant m_n$, $\mathrm{D}_t\mathrm{U}_n = (\mathrm{A}+\mathrm{B}-\mathrm{C})$ and $\mathrm{AB} = 0$. In that case, since $\mathrm{A}, \mathrm{B}, \mathrm{C} \in \{0,1\}$,

$$(\mathrm{D}_t\mathrm{U}_n)^2 = \mathrm{A}^2 + \mathrm{B}^2 + \mathrm{C}^2 - 2\mathrm{AC} - 2\mathrm{BC} = \mathrm{D}_t\mathrm{U}_n + 2\mathrm{C}(1-\mathrm{A}-\mathrm{B}).$$

Since $\mathrm{AB} = 0$,

$$\mathbf{E}\big[|(\mathrm{D}_t\mathrm{U}_n)^2 - \mathrm{D}_t\mathrm{U}_n|\big] = 2\mathbf{E}[\mathrm{C}\mathbf{1}_{\{\mathrm{A}=0\}\cap\{\mathrm{B}=0\}}] \leqslant 2(1-p_n^{m_n-1})p_n^{2m_n} \leqslant 2p_n^{2m_n}.$$

Finally,

$$A_2 \leqslant 2(n+m_n-1)(1-p_n)p_n^{2m_n}.$$

Inserting this in (5.12) and using that $(1-e^{-\lambda_n})/\lambda_n \leqslant \lambda_n^{-1}$, we get (5.11). □

5.2 Number of occurrences of a word in a DNA sequence

Identifying words with unexpected frequencies is crucial in DNA sequence analysis, and in diagnostic issues in particular. A DNA sequence can be represented by a finite series $\mathrm{X}_1\mathrm{X}_2\cdots\mathrm{X}_n$ of characters taken from the alphabet $\mathcal{A} := \{\mathrm{A}, \mathrm{C}, \mathrm{G}, \mathrm{T}\}$ where the letters stand for the four bases *adenine, cytosine, guanine* and *thymine*. Here, we model the sequence $\mathrm{X}_1\mathrm{X}_2\cdots\mathrm{X}_n$ with an homogeneous and stationary Markov chain of order m. Since a m-order Markov chain on $\mathcal{A}$ can be rewritten as a Markov chain of order 1 on $\mathcal{A}^m$, we assume hereafter that $m=1$. The invariant probability measure is denoted by μ. The aim is to compute the number of occurrences, within the sequence, of a given word W_n of size h_n (with $h_n > m$), $\mathrm{W}_n = w_1w_2\cdots w_{h_n}$. Let for any $j \in \mathcal{J}_n := \{1, \ldots, n-h_n+1\}$ the random variable

$$\mathrm{Z}_j = \mathbf{1}_{\{\mathrm{X}_j=w_j,\ldots,\mathrm{X}_{j+h_n-1}=w_{h_n}\}}.$$

Since the underlying Markov chain is homogeneous and stationary with invariant measure μ, $\mathbf{E}[\mathrm{Z}_j] = \mu(\mathrm{W}_n)$ ($j \in \mathcal{J}_n$). The number of occurrences of the word W_n is then provided by

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

We want to analyze its asymptotics when n goes to infinity and h_n grows as $\log(n)$. We propose to address this issue using the Stein-Malliavin method rather than the dependency neighbourhood study used in most related works (e.g. [26, 51, 49]). As explained in [51], the word W_n may appear in clumps; if W_n 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_n. For instance, if $\mathrm{W}_n = \mathrm{ACTAA}$, the sequence

$$\mathrm{G\underline{ACTAACTAAACTAA}TGAA\underline{ACTAA}CG}$$
$$\mathrm{ACTAACTAA}$$

has a 3-clump at position $j = 2$ and a 1-clump at position $j = 20$. Then, we must consider $(\widetilde{\mathrm{Z}}_j)_{j\in\mathcal{J}_n}$, the "declumped" sequence associated to $(\mathrm{Z}_j)_{j\in\mathcal{J}_n}$, such that $\widetilde{\mathrm{Z}}_j$ only counts occurrences that do not overlap the preceding one. As noted in the remark 2 of [51], we can define (with a slight abuse since the variables $\mathrm{X}_0, \mathrm{X}_{-1}, \dots, \mathrm{X}_{-h+2}$ may not be known) the $\widetilde{\mathrm{Z}}_j$ by

$$\widetilde{\mathrm{Z}}_j = \mathrm{Z}_j(1-\mathrm{Z}_{j-1})\cdots(1-\mathrm{Z}_{j-h_n+1})\,;\;\; j \in \mathcal{J}_n.$$

Let $\overline{\mathfrak{T}}^{(k)}(\mathrm{W}_n)$ and $\overline{\mathrm{Z}}_j^{(k)}$ $(j \in \mathcal{J}_n, k \in \mathbb{N})$ be the random variables that indicate respectively the number of k-clumps and whether there is a k-clump at position j. Let us write $\mathfrak{T}(\mathrm{W}_n) = \sum_{k\in\mathbb{N}} k\overline{\mathfrak{T}}^{(k)}(\mathrm{W}_n)$; it can be well approximated (see e.g [7, 49]) by the random variable

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

and we have $\mathrm{d_{TV}}(\mathbf{P}_{\mathfrak{T}(\mathrm{W}_n)}, \mathbf{P}_{\overline{\mathfrak{T}}(\mathrm{W}_n)}) \leqslant 2h_n\mu(\mathrm{W}_n)$. Moreover, it appears in [49] that for any $j \in \mathcal{J}_n$, $\overline{\mathrm{Z}}_j^{(k)}$ is a Bernoulli-distributed random variable with mean

$$p_k = (1-\alpha)^2\alpha^{k-1}\mu(\mathrm{W}_n) \tag{5.13}$$

where α can be written with respect to the principal periods of W_n. An explicit expression of α in the case of a first-order Markov chain is provided in [51], section 3. This last point suggests to approximate $\mathfrak{T}(\mathrm{W}_n)$ by $\overline{\mathfrak{T}}(\mathrm{W}_n)$. Let us introduce the marked binomial functional

$$\mathrm{H}_n = \sum_{j\in\mathcal{J}_n} \mathrm{V}_j \Delta\mathrm{N}_j, \tag{5.14}$$

where $(\mathrm{V}_j)_{j\in\mathbb{N}}$ is a sequence of i.i.d. random variables with geometric distribution $\mathbf{V}$ with parameter $(1-\alpha)$ and α appears in (5.13). In fact, $(1-\alpha)\alpha^{k-1}$ is the probability that the word W_n overlaps exactly k times after having occurred at position j. The sequence $(\mathrm{V}_j)_{j\in\mathbb{N}}$ is also supposed to be independent of the increments $(\Delta\mathrm{N}_j)_{j\in\mathcal{J}_n}$ of the binomial process N with intensity $(1-\alpha)\mu(\mathrm{W}_n)$ so that $\mathcal{PC}(\lambda_n, \mathbf{V})$ is exactly the Pólya-Aeppli distribution with parameters (λ_n, α) where $\lambda_n = (n-h_n+1)(1-\alpha)\mu(\mathrm{W}_n)$. Some computations highlight that $\overline{\mathfrak{T}}(\mathrm{W}_n)$ and H_n are identically distributed, so that $\mathrm{d_{TV}}(\mathbf{P}_{\overline{\mathfrak{T}}(\mathrm{W}_n)}, \mathbf{P}_{\mathrm{H}_n}) = 0$. The quantity $\mathrm{d_{TV}}\big(\mathbf{P}_{\mathrm{H}_n}, \mathcal{PC}(\lambda_n, \mathbf{V})\big)$ has yet to be controlled using Theorem 5.2, that results in the following bound.

Proposition 5.6. *Let* $\lambda_n = (n-h_n+1)(1-\alpha)\mu(\mathrm{W}_n)$. *Then,*

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

The convergence occurs since the assumption on the order of the length h_n (in $\log n$) entails that $n\mu(\mathrm{W}_n) = \mathrm{O}(1)$ (see e.g. [51]). 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" as in [51] or on the order of the magnitude of the maximal overlap as in [26].

Proof of Proposition 5.6. Consider the random variable $\mathrm{H}_n = \sum_{j\in\mathcal{J}_n} \mathrm{V}_j \Delta \mathrm{N}_j$, as defined by (5.14). Since H_n belongs to $\mathcal{H}_1$, then $\widetilde{\mathrm{L}}\mathrm{H}_n = \mathrm{H}_n$ so that $\mathrm{D}\widetilde{\mathrm{L}}^{-1}(\mathrm{H}_n - \mathbf{E}[\mathrm{H}_n]) = \mathrm{D}^+\widetilde{\mathrm{L}}^{-1}(\mathrm{H}_n - \mathbf{E}[\mathrm{H}_n]) = \mathrm{D}^+\mathrm{H}_n$. Moreover for any $(t,k) \in \mathcal{J}_n \times \mathbb{N}$, $\mathrm{D}^+_{(t,k)}\mathrm{H}_n = k$ and $\widetilde{\mathrm{D}}_{(t,k)}\mathrm{H}_n = (k-\ell)\mathbf{1}_{\{\eta(t,\ell)=1\}}$ $\mathbf{P}$-almost surely. Then the second term in (5.8) vanishes and we must still control

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

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

$$\begin{aligned}
\big|\mathbf{E}\big[\psi_{\mathrm{A}}(\mathfrak{h}(\pi_t(\eta)+&\delta_{(t,k)})) - \psi_{\mathrm{A}}(\mathfrak{h}(\eta)+k)\big]\big| \\
&= \Big|\sum_{\ell\in\mathbb{N}} \mathbf{E}\Big[\Big(\psi_{\mathrm{A}}\big(\mathfrak{h}(\pi_t(\eta)+\delta_{(t,k)})\big) - \psi_{\mathrm{A}}\big(\mathfrak{h}(\eta)+k\big)\Big)\mathbf{1}_{\{\eta(t,\ell)=1\}}\Big]\Big| \\
&= \Big|\sum_{\ell\in\mathbb{N}} \mathbf{E}\Big[\Big(\psi_{\mathrm{A}}\big(\mathfrak{h}(\pi_t(\eta)+\delta_{(t,k)})\big) - \psi_{\mathrm{A}}\big(\mathfrak{h}(\pi_t(\eta)+\delta_{(t,\ell)})+k\big)\Big)\mathbf{1}_{\{\eta(t,\ell)=1\}}\Big]\Big| \\
&\leqslant \|\nabla\psi_{\mathrm{A}}\|_\infty \sum_{\ell\in\mathbb{N}} \mathbf{E}\Big[\big|\mathrm{H}_n^{\neg t}+k - \big(\mathrm{H}_n^{\neg t}+k+\ell\big)\big|\mathbf{1}_{\{\eta(t,\ell)=1\}}\Big] \\
&= \|\nabla\psi_{\mathrm{A}}\|_\infty \sum_{\ell\in\mathbb{N}} \ell\mu(\mathrm{W}_n)(1-\alpha)^2\alpha^{\ell-1} = \|\nabla\psi_{\mathrm{A}}\|_\infty\mu(\mathrm{W}_n).
\end{aligned}$$

Then,

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

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

6 Application 2: optimal hedging in the trinomial model

Trinomial model In a filtered probability space $(\Omega^{\mathrm{tri}}, \mathcal{A}^{\mathrm{tri}}, (\mathcal{F}_t^{\mathrm{tri}})_{t\in\mathbb{T}}, \mathbf{P}^{\mathrm{tri}})$, we consider a simple financial market modelled by two assets, i.e., a couple of $\mathbb{R}_+$-valued processes $(\mathrm{A}_t^{\mathrm{tri}}, \mathrm{S}_t^{\mathrm{tri}})_{t\in\mathbb{T}}$, and $\mathbb{T} = \mathbb{Z}_+ \cap [0,T]$ ($T\in\mathbb{N}$) is called the trading interval. Denote also $\mathbb{T}^* = \mathbb{T}\setminus\{0\}$. The *riskless asset* $(\mathrm{A}_t^{\mathrm{tri}})_{t\in\mathbb{T}}$, deterministic, is defined by $\mathrm{A}_t^{\mathrm{tri}} = a_0(1+r)^t$ ($t\in\mathbb{T}$) for $r\in\mathbb{R}_+^*$; w.l.o.g. we take $a_0 = 1$ and $r<1$. The stock price $(\mathrm{S}_t^{\mathrm{tri}})_{t\in\mathbb{T}}$ which models the *risky asset* is the $\mathcal{F}$-adapted process such that $\mathrm{S}_0^{\mathrm{tri}} = 1$ and for $t\in\mathbb{T}^*$,

$$\Delta\mathrm{S}_t^{\mathrm{tri}} = \theta_t^{\mathrm{tri}}\mathrm{S}_{t-1}^{\mathrm{tri}}, \tag{6.1}$$

where $\theta_t^{\mathrm{tri}} := b\mathbf{1}_{\{\mathrm{X}_t^{\mathrm{tri}}=1\}} + a\mathbf{1}_{\{\mathrm{X}_t^{\mathrm{tri}}=-1\}} + r\mathbf{1}_{\{\mathrm{X}_t^{\mathrm{tri}}=0\}}$ and the real numbers a,r,b satisfy $-1<a<0<r<b$. The family $\{\mathrm{X}_t^{\mathrm{tri}},\ t\in\mathbb{T}^*\}$ consists of i.i.d. $\{-1,0,1\}$-valued random variables such that $\mathbf{P}(\{\mathrm{X}_1^{\mathrm{tri}}=k\}) =: p_k^{\mathrm{tri}}$ for all $k\in\{-1,0,1\}$.

6.1 A surrogate model: the jump-binomial model

Let us replace the trinomial model by a more computationally amenable one called *jump-binomial model* (J-Bi) and that is based on a MBP. In $(\Omega,\mathcal{A},(\mathcal{F}_t)_{t\in\mathbb{T}},\mathbf{P})$, let us define the riskless asset $(\mathrm{A}_t)_{t\in[\![0,T]\!]}$ by $\mathrm{A}_0 = 1$ and $\mathrm{A}_t = \mathrm{A}_0(1+r)^t$ for all $t\in[\![1,T]\!]$. The stock price is the $\mathcal{F}$-adapted process $(\mathrm{S}_t)_{t\in[\![0,T]\!]}$ with (deterministic) initial value $\mathrm{S}_0 = 1$ that satisfies for any $t\in[\![1,T]\!]$,

$$\Delta\mathrm{S}_t = \theta_t\,\mathrm{S}_{t-1}, \tag{6.2}$$

with $\theta_t := r\mathbf{1}_{\{\eta(t,\cdot)=0\}} + b\mathbf{1}_{\{\eta(t,1)=1\}} + a\mathbf{1}_{\{\eta(t,-1)=1\}}$.

There exists a *correspondence* between the classical trinomial model and our jump-binomial model: the role played by the random variables $\mathrm{X}_t^{\mathrm{tri}}$ in the classical trinomial model is held here (in the jump-binomial model) by the i.i.d random variables $\Delta \mathrm{Y}_t := \eta(t,1) - \eta(t,-1)$. This correspondence can be informally illustrated through the following figures.

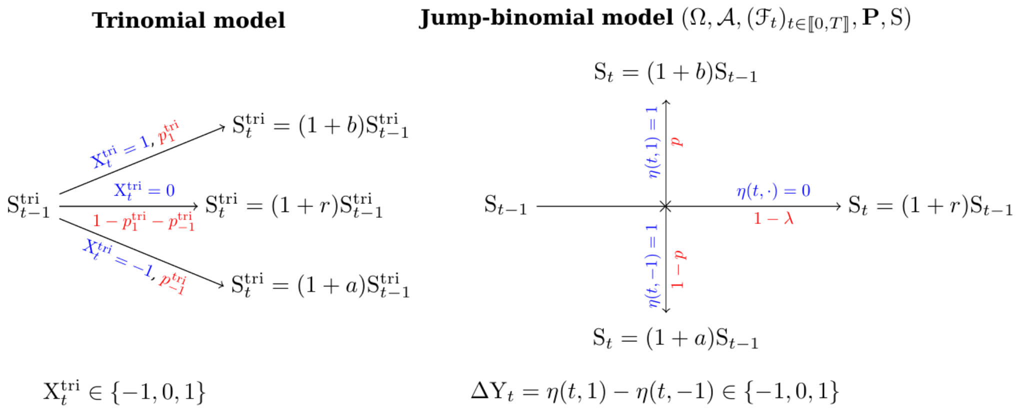

Figure 2: Trinomial model VS Jump-binomial model

Equivalence in law of the ratios By setting $\mathrm{S}_0 = \mathrm{S}_0^{\mathrm{tri}}$, $\lambda = p_{-1}^{\mathrm{tri}} + p_1^{\mathrm{tri}} > 0$ and $p = p_1^{\mathrm{tri}}/\lambda$, we get for all $s \in \mathbb{R}_+^*$,

$$\mathbf{E}\left[s^{\mathrm{S}_t/\mathrm{S}_{t-1}}\right] = \mathbf{E}\left[s^{1+\theta_t}\right] = s^{1+r}(1-\lambda) + s^{1+b}\,\lambda p + s^{1+a}\,\lambda(1-p) = \mathbf{E}_{\mathbf{P}^{\mathrm{tri}}}\left[s^{\mathrm{S}_t^{\mathrm{tri}}/\mathrm{S}_{t-1}^{\mathrm{tri}}}\right]. \quad (6.3)$$

All the results that can be stated in the jump-binomial models can be retrieved in the trinomial model defined by (6.1) by virtue of the identity (6.3) and the correspondence (see Figure 2) between the two models. The introduction of the jump-binomial model has been in fact motivated by the following remark. A Karatzas-Ocone-type hedging formula for replicable claims in the trinomial model (underlying by a sequence of $\{-1,0,1\}$-valued i.i.d. variables) can not be derived from the Clark-Ocone formula stated in [15] (Theorem 3.3.) because of the $\mathcal{F}_k$-measurability of the term $\mathrm{D}_k\mathbf{E}[\mathrm{F}|\mathcal{F}_k]$. On the contrary, the Clark formula (4.3) for MBP writes as a stochastic integral with a predictable integrand, from which we will obtain the replication strategy.

Incompleteness As explained in [50], the trinomial model stands for an incomplete market; so does the jump-binomial model. Adapting their arguments in our context, we can show that the set of probability measures $\mathbf{M}$ equivalent to $\mathbf{P}$ and with respect to which $(\overline{\mathrm{S}}_t)_{t\in\mathbb{T}}$ is a $(\mathbf{M}, \mathcal{F})$-martingale coincides with the interior of a convex polyhedron with 2^T vertices.

Prospects The jump-binomial model can advantageously replace the trinomial for other purposes such that utility maximization problems. In [27], the author recently obtained explicit expressions of the optimal expected (logarithmic, exponential and power) utility of an insider whose level of information is given by an initially enlarged filtration.

6.2 Loss minimizing hedging in the trinomial/jump-binomial model

Portfolio management in incomplete markets The value of the portfolio at $t \in \mathbb{T}$ is

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

where $(\alpha_t, \varphi_t)_{t\in\mathbb{T}}$ is a couple of $\mathcal{F}$-predictable processes modelling respectively the amounts of riskless and risky assets held in the portfolio. Its discounted value is $\overline{\mathrm{V}}_t := \mathrm{V}_t/\mathrm{A}_t$. We aim at proving a hedging formula; That is, given a nonnegative $\mathcal{F}_T$-measurable random variable F (called *claim*), to find an *admissible* strategy $\psi = (\alpha, \varphi)$ that is *self-financing*, i.e., such that

$$\mathrm{A}_t\,(\alpha_{t+1} - \alpha_t) + \mathrm{S}_t\,(\varphi_{t+1} - \varphi_t) = 0\ ;\ \text{for}\, t \in \mathbb{T} \setminus \{T\}, \tag{6.4}$$

and whose corresponding portfolio value satisfies $\mathrm{V}_0 > 0,\ \mathrm{V}_t \geqslant 0$ for all $t \in \mathbb{T} \setminus \{T\}$, and $\mathrm{V}_T = \mathrm{F}$. In an incomplete market, there is no systematic hedging formula, since all claims are not attainable: They have an *intrinsic risk* (see [52]) so that one aims at reducing the *a priori* risk to this minimal component (see [24, 52, 54]). The question of hedging in an incomplete market has been widely investigated for years (e.g. in [14, 23] in continuous time, in [54] in discrete time). As the jump-binomial model is incomplete, we study the optimization problem in return:

$$\min_{\psi\in\mathcal{S}} \mathbf{E}\big[(\mathrm{F} - x - \overline{\mathrm{V}}_T(\psi))^2\big], \tag{6.5}$$

where the claim F and the initial capital $x \in \mathbb{R}_+^*$ are given, and $\mathcal{S}$ is the set of $\mathcal{F}$-predictable admissible strategies. The *mean-variance tradeoff* process $(\mathrm{K}_t)_{t\in\mathbb{T}^*}$ (defined in [54], (0.2)) is given for any $t \in \mathbb{T}^*$ by

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

We also introduce the discrete analogue of the *minimal martingale measure* (see [23]), i.e., the signed measure $\widehat{\mathbf{P}}$ defined on $(\Omega, \mathcal{F})$ such that

$$\frac{\mathrm{d}\widehat{\mathbf{P}}}{\mathrm{d}\mathbf{P}} = \prod_{t=1}^{\mathrm{T}} \frac{1 - \theta_t \Delta\overline{\mathrm{S}}_t}{1 - \theta_t \mathbf{E}\big[\Delta\overline{\mathrm{S}}_t | \mathcal{F}_{t-1}\big]}, \tag{6.6}$$

where $(\theta_t)_{t\in\mathbb{T}^*}$ is the $\mathcal{F}$-predictable process such that $\theta_t = \mathbf{E}[\Delta\overline{\mathrm{S}}_t \,|\, \mathcal{F}_{t-1}\big]/\mathbf{E}[(\Delta\overline{\mathrm{S}}_t)^2 \,|\, \mathcal{F}_{t-1}\big]$, for $t \in \mathbb{T}^*$. Last, consider the Kunita-Watanabe decomposition of F (see e.g. [34, 54]), i.e., the unique couple of processes $(\xi^{\mathrm{F}}, \mathrm{L}^{\mathrm{F}})$ where ξ^{F} is a square-integrable admissible strategy and L^{F} is a $\mathcal{F}$-martingale, strongly orthogonal to S, with null initial value and such that

$$\mathrm{F} = \mathrm{F}_0 + \sum_{t\in\mathbb{T}^*} \xi_t^{\mathrm{F}} \Delta\overline{\mathrm{S}}_t + \mathrm{L}_T^{\mathrm{F}}\ ;\ \mathbf{P}\text{−a.s.}.$$

An expression of the quadratic-loss minimizing strategy is given in [53] (Proposition 4.3).

Theorem 6.1 (Schweizer, 1992)**.** *Provided* $(\mathrm{K}_t)_{t\in\mathbb{T}^*}$ *is deterministic, the solution of* (6.5) *is given by*

$$\varphi_t^* = \xi_t^{\mathrm{F}} + \frac{\mathbf{E}\big[\Delta\overline{\mathrm{S}}_t \,|\, \mathcal{F}_{t-1}\big]}{\mathbf{E}\big[(\Delta\overline{\mathrm{S}}_t)^2 \,|\, \mathcal{F}_{t-1}\big]} (\widehat{\mathbf{E}}\big[\mathrm{F}|\mathcal{F}_t\big] - x - \overline{\mathrm{V}}_{t-1}(\varphi^*)) \tag{6.7}$$

where $\widehat{\mathbf{E}}$ *denotes the expectation with respect to the measure* $\widehat{\mathbf{P}}$*, i.e., the minimal martingale measure defined by* (6.6)*.*

If the claim F is attainable, then $\varphi^* = \xi^{\mathrm{F}}$. The term ξ^{F} in (6.7) can be interpreted as a pure hedging demand, and the second one as a demand for mean-variance purposes (see [53]).

Loss minimizing hedging in jump-binomial model We can now solve (6.5) in (J-Bi).

Lemma 6.2. *The mean-variance tradeoff process of the ternary model is deterministic.*

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

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

is a deterministic constant. Hence the result. □

Letting $\rho := \lambda q/(1-\lambda p)$, the family $\mathcal{R}$ provided by Gram-Schmidt process (3.2) is such that

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

Lemma 6.3 (Kunita-Watanabe decomposition in the jump-binomial model)**.** *For any claim* $\mathrm{F} \in \mathrm{L}^2(\mathbf{P})$ *there exist a square-integrable admissible strategy* ξ^{F} *and a* $\mathcal{F}$*-martingale* L^{F}*, strongly orthogonal to* $\overline{\mathrm{S}}$*, with null initial value such that*

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

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

$$\xi_t^{\mathrm{F}} = \frac{1}{\overline{\mathrm{S}}_{t-1}}\Big(\sum_{k\in\mathrm{E}} w_{t,k}\widehat{\mathbf{E}}\left[\mathrm{D}_{(t,k)}\mathrm{F}|\mathcal{F}_{t-1}\right]\Big) \;\; \textit{and} \;\; \mathrm{L}_t^{\mathrm{F}} = \mathbf{E}\Big[\mathrm{F} - \sum_{s\in\mathbb{T}^*} \xi_s^{\mathrm{F}}\Delta\overline{\mathrm{S}}_s \,\Big|\, \mathcal{F}_t\Big] - \mathbf{E}\Big[\mathrm{F} - \sum_{s\in\mathbb{T}^*}\xi_s^{\mathrm{F}}\Delta\overline{\mathrm{S}}_s\Big], \tag{6.8}$$

where $\mathbf{E}[\mathrm{L}_0^{\mathrm{F}}] = 0$*, the sequence* $w = (w_{t,k})_{(t,k)\in\mathbb{X}}$ *is defined by*

$$w_{t,1} = \frac{(b-a\rho)\kappa_1}{(b-a\rho)^2\kappa_1 + a^2\kappa_{-1}}, \quad w_{t,-1} = \frac{a\kappa_{-1}}{(b-a\rho)^2\kappa_1 + a^2\kappa_{-1}}.$$

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

$$\frac{\mathrm{d}\widehat{\mathbf{P}}}{\mathrm{d}\mathbf{P}} = \prod_{t\in\mathbb{T}} \frac{1-\theta_t\Delta\overline{\mathrm{S}}_t}{1-\theta_t\mathbf{E}[\Delta\overline{\mathrm{S}}_t|\mathcal{F}_{t-1}]}, \tag{6.9}$$

with for $t \in \mathbb{T}^*$*,*

$$\theta_t = \frac{\overline{\mathrm{S}}_{t-1}(\lambda(bp+aq)-r)}{\overline{\mathrm{S}}_{t-1}^2(\lambda^2(b^2p+a^2q)+r^2-2\lambda(bp+aq))} = \frac{\lambda(bp+aq)-r}{\overline{\mathrm{S}}_{t-1}((b-a\rho)^2\kappa_1 + a^2\kappa_{-1})}.$$

Proof. From the proof of Lemma 2.7 in [54], ξ_t^{F} can be simply written

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

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

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

where we have used that $\mathbf{E}[\Delta\mathrm{R}_{(t,\ell)}\Delta\mathrm{R}_{(t,k)}|\mathcal{F}_{t-1}] = 0$ for $\ell \neq k$ due to the independence of $\Delta\mathrm{R}_{(t,\cdot)}$ and $\mathcal{F}_{t-1}$ and the orthogonality of the family $\mathcal{R}$. The sequence $v = (v_t)_{t\in\mathbb{T}^*}$

is defined by $v_t = \lambda^2(b^2p + a^2q) + r^2 - 2\lambda(bp + aq) = (b - a\rho)^2\kappa_1 + a^2\kappa_{-1}$, by using that $b\Delta\mathrm{Z}_{(t,1)} + a\Delta\mathrm{Z}_{(t,-1)} = (b - a\rho)\Delta\mathrm{R}_{(t,1)} + a\Delta\mathrm{R}_{(t,-1)}$. The sequence $u = (u_{t,k})_{(t,k)\in\mathbb{X}}$ is given by

$$u_{t,1} = (b - a\rho)\kappa_1 \quad \text{and} \quad u_{t,-1} = a\kappa_{-1}.$$

Last, let $w_{t,k} = u_{t,k}/v_t$ for $(t,k) \in \mathbb{X}$. The proof is complete. □

The two previous lemmas directly lead to the main result of the section.

Theorem 6.4 (Loss quadratic minimizing strategy). *Let $\widehat{\mathbf{P}}$ be the minimal martingale measure defined by* (6.9) *and let a claim* F. *The quadratic loss minimizing hedge φ^* is given by*

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

where for any $t \in \mathbb{T}^*$, ξ^{F}_t is given by The Kunita-Watanabe decomposition (6.8). *Moreover, the quota of the riskless asset $(\mathrm{A}_t)_{t\in\mathbb{T}}$ is given by $\alpha_0 = \widetilde{\mathbf{E}}[\mathrm{F}]/\mathrm{S}_0$ and for any $t \in \mathbb{T}^*$,*

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

Proof. Since the mean-variance process is deterministic by Lemma 6.2, it is enough to incorporate the result of Lemma 6.3 into Theorem 6.1. The process $(\alpha_t)_{t\in\mathbb{T}}$ is defined by the self-financing condition (6.4). □

The expression (6.8) of ξ^{F} that is the replication strategy if F is attainable, reminds that obtained in the binomial model (see [46], Proposition 1.14.4) or in generalized discrete-time complete market models that can be constructed from an obtuse random walk as in [25].

A Technical results

Proof of Proposition 3.10. We will use the following result (whose proof is similar to [46], Proposition 6.2.5): for $g \in \mathrm{L}^2(\nu)$ and $f_n \in \mathrm{L}^2(\nu^{\circ n})$,

$$\mathrm{J}_{n+1}(g \circ f_n) = n \sum_{(t,k)\in\mathbb{X}} \mathrm{J}_n\big(f_n(\star,(t,k)) \circ g(\cdot)\mathbf{1}_{[t-1]^n}(\star,\cdot)\big)\Delta\mathrm{R}_{(t,k)} + \sum_{(t,k)\in\mathbb{X}} g(t,k)\mathrm{J}_n(f_n\mathbf{1}_{[t-1]^n})\Delta\mathrm{R}_{(t,k)}, \quad \text{(A.1)}$$

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

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

For any $T \in \mathbb{N}, t \in [T-1]$ define

$$\zeta^T_t = 1 + \sum_{n=1}^{T} \frac{1}{n!}\mathrm{J}_n(h^{\otimes n}\mathbf{1}_{[t]^n}),$$

where we assume w.l.o.g. that $\mathbf{E}[\zeta_t^T]=1$. Let T be such that $T>t$. Then,

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

where we used (A.1) in the second line and (3.7) in the penultimate one. Since by (3.16), for all $t\in\mathbb{N}$, ζ_t^T tends to $\xi_t(h)$ $\mathbf{P}$-almost surely when T goes to 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 $(\xi_t(h))_{t\geqslant 1}$ satisfies the equation in differences

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

where $\mathrm{J}_1(h)=\mathrm{J}_1(g\,;\mathcal{Z})$. Provided the product converges, define the sequence of exponential products $(\xi_t^{\mathcal{Z}}(g))_{t\geqslant 1}$ – that stand for the Doléans exponentials with respect to $\mathcal{Z}$ – by

$$\begin{aligned}
\xi_t(h)&=\xi_t^{\mathcal{Z}}(g)\\
&=\prod_{s=1}^{t}\Big(1+\sum_{k\in \mathrm{E}} g(s,k)\big(\mathbf{1}_{\{\eta(s,k)=1\}}-\lambda\mathbf{Q}(\{k\})\big)\Big)=1+\sum_{s=1}^{t}\Big(\sum_{k\in \mathrm{E}} h(s,k)\Delta \mathrm{R}_{(s,k)}\Big)\xi_{s-1}^{\mathcal{Z}}(g).
\end{aligned}$$

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

Proof of Lemma 4.2. It is enough to state the result for $\mathrm{F}=\xi(h)$, with $h\in \mathrm{L}^2(\nu)$. By (3.17),

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

It follows from (4.1) that for any $(\mathbf{t}_n,\mathbf{k}_n)\in\mathbb{X}^n$ and any set $\big\{(s_i,\ell_i),\, i\in \mathrm{J}, |\mathrm{J}|=m\big\}$ with $m>n$, there exists $i_0\in \mathrm{J}$ such that $(s_{i_0},\ell_{i_0})\notin(\mathbf{t}_n,\mathbf{k}_n)$. Then, by independence of the $\Delta \mathrm{R}_{(t,\cdot)}$,

$$\mathbf{E}\Big[\mathrm{D}^{(n)}_{(\mathbf{t}_n,\mathbf{k}_n)}\Big(\prod_{i\in \mathrm{J}} h(s_i,\ell_i)\Delta \mathrm{R}_{(s_i,\ell_i)}\Big)\Big]$$

$$= \mathbf{E}\Big[\Delta \mathrm{R}_{(s_{i_0},\ell_{i_0})}\Big]\mathbf{E}\Big[\mathrm{D}^{(n)}_{(\mathbf{t}_n,\mathbf{k}_n)}\Big(\prod_{i\in \mathrm{J}\setminus\{i_0\}} h(s_i,\ell_i)\Delta \mathrm{R}_{(s_i,\ell_i)}\Big)\Big] = 0.$$

For $(t,k)\in\mathbb{X}$ let $\mathfrak{r}_{(t,k)}$ be the representative of $\Delta\mathrm{R}_{(t,k)}$. With a similar argument we prove the same result for any set $\big\{(s_i,\ell_i),\, i\in \mathrm{J}, |\mathrm{J}|=n\big\}$ different from $(\mathbf{t}_n,\mathbf{k}_n)$ so that the expectation does not vanish if and only if $\big\{(s_i,\ell_i),\, i\in \mathrm{J}, |\mathrm{J}|=n\big\} = (\mathbf{t}_n,\mathbf{k}_n) = \big\{(t_i,k_i),\, i\in [n]\big\}$. In that case,

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

Besides, using the alternative characterization of $\mathrm{F}=\xi(h)$, and the orthogonality of the centered variables $\Delta\mathrm{R}$,

$$\begin{aligned}
\mathbf{E}\Big[\mathrm{F}\prod_{i=1}^{n}\frac{1}{\kappa_i}\Delta\mathrm{R}_{(t_i,k_i)}\Big] &= \mathbf{E}\Big[\mathbf{E}[\xi(h)]\prod_{i=1}^{n}\frac{1}{\kappa_i}\Delta\mathrm{R}_{(t_i,k_i)}\Big]\\
&\quad + \mathbf{E}\Big[\prod_{s\in\mathbb{N}}\Big(1+\sum_{k\in\mathrm{E}} h(s,k)\Delta\mathrm{R}_{(s,k)}\Big)\prod_{i=1}^{n}\frac{1}{\kappa_i}\Delta\mathrm{R}_{(t_i,k_i)}\Big]\\
&= \mathbf{E}\Big[\prod_{i=1}^{n}\Big(1+\sum_{k\in\mathrm{E}} h(t_i,k)\Delta\mathrm{R}_{(t_i,k)}\Big)\frac{1}{\kappa_i}\Delta\mathrm{R}_{(t_i,k_i)}\Big] = \prod_{i=1}^{n} h(t_i,k_i).
\end{aligned}$$

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

Proof of Lemma 4.3. It is enough to state the equality for $\mathrm{F}=\xi(f)$, $\mathrm{G}=\xi(g)$ where $f,g\in\mathrm{L}^2(\nu)$ and such that $\mathbf{E}[\mathrm{F}]=\mathbf{E}[\mathrm{G}]=1$. On the one hand, there exists $T\in\mathbb{N}$ such that

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

On the other hand, for any $n\in[T]$ and $\mathrm{I}_n\subset[T]$ of cardinality n,

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

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

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

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

B Bounds based on contractions and application

Frame : Simple binomial process In this part, assume that $\mathrm{E} = \{1\}$. Then, $\mathbb{X}$ can be identified to $\mathbb{N}$ and $\nu(\mathrm{A}) = \lambda|\mathrm{A}|$ for $\mathrm{A} \in \mathcal{P}(\mathbb{N})$. We have $\mathrm{L}^2(\mathbf{P}) = \bigoplus_{n\in\mathbb{Z}_+} \mathcal{H}_n$ where $\mathcal{H}_n = \{\mathrm{J}_n(f_n)\ ;\ f_n \in \mathrm{L}^2(\nu^{\circ n})\}$ and

$$\mathrm{J}_n(f_n) = \sum_{\mathbf{t}_n\in\mathbb{N}^{n,<}} f_n(t_1,\dots,t_n)\prod_{i=1}^n \Delta\mathrm{Z}_{t_i} \quad \text{with } \Delta\mathrm{Z}_{t_i} = \mathbf{1}_{\{\eta(t_i)=1\}} - \lambda.$$

Star contractions Let $k, q \in \mathbb{N}$. For symmetric functions $f \in \mathrm{L}^2(\nu^{\circ k})$ and $g \in \mathrm{L}^2(\nu^{\circ q})$, the contraction kernel on $\mathbb{N}^{k+q-r-\ell}$ and denoted $f \star_r^\ell g$ is defined by

$$f \star_r^\ell g(\mathbf{y}_\ell, \mathbf{z}_{r-\ell}, \mathbf{t}_{k-r}, \mathbf{s}_{q-r}) = \int_{\mathbb{N}^\ell} f(\mathbf{y}_\ell, \mathbf{z}_{r-\ell}, \mathbf{t}_{k-r}) \times g(\mathbf{y}_\ell, \mathbf{z}_{r-\ell}, \mathbf{s}_{q-r})\mathrm{d}\nu^\ell(\mathbf{y}_\ell) \tag{B.1}$$

where $(\mathbf{y}_\ell, \mathbf{z}_{r-\ell}, \mathbf{t}_{k-r}, \mathbf{s}_{q-r}) = (y_1,\dots,y_\ell, z_1, z_{r-\ell}, t_1,\dots,t_{k-r}, s_1,\dots,s_{q-r})$.

A general bound in terms of contractions Consider a random variable $\mathrm{F} = \mathbf{E}[\mathrm{F}] + \sum_{n=1}^m \mathrm{J}_n(f_n) =: \mathbf{E}[\mathrm{F}] + \sum_{n=1}^m \mathrm{F}_n$ for some $m \in \mathbb{N}$. Using the same approach as in [29] we can show that

$$\begin{aligned}
&\sqrt{\mathbf{E}\Big[\Big|\mathrm{var}[\mathrm{F}] - \int_{\mathbb{N}} \widetilde{\mathrm{D}}_t\mathrm{F}(-\mathrm{D}_t\mathrm{L}^{-1}\mathrm{F})\mathrm{d}\nu(t)\Big|^2\Big]}\\
&\quad\leqslant \sum_{i\in[m_n]} \sqrt{\mathbf{E}\Big[\Big|\mathrm{var}[\mathrm{F}_i] - \int_{\mathbb{N}} \widetilde{\mathrm{D}}_t\mathrm{F}_i(-\mathrm{D}_t\mathrm{L}^{-1}\mathrm{F}_i)\mathrm{d}\nu(t)\Big|^2\Big]}\\
&\qquad + \sum_{(i,j)\in[m_n]^{\neq}} \sqrt{\mathbf{E}\Big[\Big|\int_{\mathbb{N}} \widetilde{\mathrm{D}}_t\mathrm{F}_i(-\mathrm{D}_t\mathrm{L}^{-1}\mathrm{F}_j)\mathrm{d}\nu(t)\Big|\Big]}.
\end{aligned}$$

Adapting the result of Proposition 5.5 in [44], for any $i, j \in [m]$ $(i < j)$ we can prove the existence of constants $\mathfrak{c}_i, \mathfrak{c}_{ij}$ such that

$$\mathbf{E}\Big[\Big|\mathrm{var}[\mathrm{F}_i] - \int_{\mathbb{N}} \widetilde{\mathrm{D}}_t\mathrm{F}_i(-\mathrm{D}_t\mathrm{L}^{-1}\mathrm{F}_i)\mathrm{d}\nu(t)\Big|^2\Big] \leqslant \mathfrak{c}_i \max_{(r,\ell)\in[i]\times[r\wedge i-1]} \|f_i \star_r^\ell f_i\|_{\mathrm{L}^2(\nu^{2i-r-\ell})},$$

and

$$\mathbf{E}\Big[\Big|\int_{\mathbb{N}} \widetilde{\mathrm{D}}_t\mathrm{F}_i(-\mathrm{D}_t\mathrm{L}^{-1}\mathrm{F}_j)\mathrm{d}\nu(t)\Big|^2\Big] \leqslant \mathfrak{c}_{ij} \max_{(r,\ell)\in[i]\times[r]} \|f_i \star_r^\ell f_j\|_{\mathrm{L}^2(\nu^{i+j-r-\ell})}.$$

Then, there exists a constant $\mathfrak{c}$ depending on the $\mathfrak{c}_i$'s and the $\mathfrak{c}_{ij}$'s such that

$$\begin{aligned}
&\sqrt{\mathbf{E}\Big[\Big|\mathrm{var}[\mathrm{F}] - \int_{\mathbb{N}} \widetilde{\mathrm{D}}_t\mathrm{F}(-\mathrm{D}_t\mathrm{L}^{-1}\mathrm{F})\mathrm{d}\nu(t)\Big|^2\Big]}\\
&\quad\leqslant \mathfrak{c}\Big[\max_{(i,r,\ell)\in[m_n]\times[i]\times[r\wedge i-1]} \|f_i \star_r^\ell f_i\|_{\mathrm{L}^2(\nu^{2i-r-\ell})}\\
&\qquad + \max_{(i,j,r,\ell)\in[m_n]\times[\![i+1,m_n]\!]\times[i]\times[r]} \|f_i \star_r^\ell f_j\|_{\mathrm{L}^2(\nu^{i+j-r-\ell})}\Big] =: \mathfrak{c}\big(\mathfrak{m}_1^\star + \mathfrak{m}_2^\star\big).
\end{aligned} \tag{B.2}$$

Application to head run problem (section 5.1)

Proof of Lemma 5.5. The proof is based on the following decomposition.

Lemma B.1. *The chaos decomposition of* U_n *writes*

$$\mathrm{U}_n = \mathbf{E}[\mathrm{U}_n] + \sum_{k=1}^{m_n} \mathrm{J}_k(u_k)$$

where for any $k \in [m_n]$, $\mathbf{t}_k = (t_1, \ldots, t_k) \in [n-1]^<$,

$$u_k(\mathbf{t}_k) = \frac{1}{k!}\Big[p_n^{m_n-k}\prod_{i=1}^{k}\mathbf{1}_{[m_n]}(t_j) + (t_1+m_n-k)(1-p_n)\mathbf{1}_{\{|t_1-t_k|\leqslant m_n\}}p_n^{m_n-k} - \mathbf{1}_{\{|t_1-t_k|\leqslant m_n\}}p_n^{m_n-k+1}\prod_{i=1}^{k}\mathbf{1}_{[n-1]}(t_j)\Big].$$

Proof. As $\mathrm{D}^{(2)}_{t,s} = \mathrm{D}^{(2)}_{s,t}$ we can assume that the t_i are ordered. Let us proceed by induction on k. For $k = 1$, follows from (5.10) and (4.2) that

$$u_1(t) = \mathbf{E}[\mathrm{D}_t\mathrm{U}_n] = \mathbf{1}_{[m_n]}(t)p_n^{m_n-1} + (1-p_n)p_n^{m_n-1}(\min(m_n,t)-1) - \mathbf{1}_{[n-1]}(t)p_n^{m_n}.$$

Assume that there exists $k \in [m_n]$ such that

$$\mathrm{D}^{(k-1)}_{\mathbf{t}_{k-1}}\mathrm{U}_n = \prod_{j=1}^{k-1}\mathbf{1}_{[m_n]}(t_j)\prod_{i=1, i\notin\mathbf{t}_{k-1}}^{m_n}\Delta\mathrm{N}^n_i + \sum_{i=\max(t_{k-1}-m_n,1)}^{t_1-1}(1-\Delta\mathrm{N}^n_i)\prod_{\ell=1, i+\ell\notin\mathbf{t}_{k-1}}^{m_n}\Delta\mathrm{N}^n_{i+\ell} - \prod_{j=1}^{k-1}\mathbf{1}_{[t_1+m_n]}(t_j)\mathbf{1}_{[n-1]}(t_j)\prod_{\ell=1}^{m_n}\Delta\mathrm{N}^n_{t_1+\ell}.$$

Since we have assumed that $t_k > t_{k-1}$ the latter equation still holds at rank k by replacing whenever necessary $k-1$ by k. We deduce from Stroock's formula that

$$\begin{aligned} u_k(\mathbf{t}_k) &= \frac{\mathbf{E}\big[\mathrm{D}^{(k)}_{\mathbf{t}_k}\mathrm{U}_n\big]}{k!} \\ &= \frac{1}{k!}\Big[\prod_{j=1}^{k}\mathbf{1}_{[m_n]}(t_j)p_n^k + \sum_{i=\max(t_k-m_n,1)}^{t_1-1}(1-p_n)p_n^{m-k} - \prod_{j=1}^{k}\mathbf{1}_{[n-1]}(t_j)\mathbf{1}_{[t_1+m_n]}(t_j)p_n^{m_n}\Big] \\ &= \frac{1}{k!}\Big[\prod_{j=1}^{k}\mathbf{1}_{[m_n]}(t_j)p_n^k \\ &\quad + \min(t_1-t_k+m_n,t_1)(1-p_n)p_n^{m_n-k} - \prod_{j=1}^{k}\mathbf{1}_{[n-1]}(t_j)\mathbf{1}_{[t_1+m_n]}(t_j)p_n^{m_n}\Big], \end{aligned}$$

which achieves the proof. ☐

Let us go back to the proof of Lemma 5.5. Assume first for any $k, q \leqslant m_n$ such that $q > p$ there exists one element in $(\mathbf{y}_\ell, \mathbf{z}_{r-\ell}, \mathbf{t}_{k-r}, \mathbf{s}_{q-r}) = (y_1, \ldots, y_\ell, z_1, z_{r-\ell}, t_1, \ldots, t_{k-r}, s_1, \ldots, s_{q-r})$ larger than m_n. Let $\overline{\mathbf{y}}^{\mathbf{t}} := \max_{i,j,l}(y_i, z_j, t_l)$, $\underline{\mathbf{y}}_{\mathbf{t}} := \min_{i,j,l}(y_i, z_j, s_l)$, $\overline{\mathbf{y}}^{\mathbf{s}} :=$

$\max_{i,j,l}(y_i,z_j,s_l)$ and $\underline{\mathbf{y}}_{\mathbf{s}} := \min_{i,j,l}(y_i,z_j,s_l)$. We can show that:

$$\begin{aligned}
&u_k \star_r^\ell u_q(\mathbf{z}_{r-\ell},\mathbf{t}_{p-r},\mathbf{s}_{q-r})\\
&=\frac{1}{k!q!}\sum_{\mathbf{y}_\ell}\min(\underline{\mathbf{y}}_{\mathbf{t}}-\overline{\mathbf{y}}^{\mathbf{t}}+m_n,\underline{\mathbf{y}}_{\mathbf{t}})\min(\underline{\mathbf{y}}_{\mathbf{s}}-\overline{\mathbf{y}}^{\mathbf{s}}+m_n,\underline{\mathbf{y}}_{\mathbf{s}})(1-p_n)^2p_n^{2m_n-k-q_n+\ell}+\mathrm{o}(p_n^{2m_n-q+\ell})\\
&=\frac{1}{k!q!}\sum_{\substack{\mathbf{y}_\ell/\overline{\mathbf{y}}^{\mathbf{t}}<\underline{\mathbf{y}}_{\mathbf{t}}+m_n\\ \overline{\mathbf{y}}^{\mathbf{s}}<\underline{\mathbf{y}}_{\mathbf{s}}+m_n}}(\underline{\mathbf{y}}_{\mathbf{t}}-\overline{\mathbf{y}}^{\mathbf{t}}+m_n)(\underline{\mathbf{y}}_{\mathbf{s}}-\overline{\mathbf{y}}^{\mathbf{s}}+m_n)(1-p_n)^2p_n^{2m_n-k-q+\ell}+\mathrm{o}(p_n^{2m_n-q+\ell})\\
&\leqslant\frac{1}{k!q!}(1-p_n)^2p_n^{2m_n-k-q+\ell}\sum_{i=\ell}^{m_n-1}\binom{i}{\ell}^2,
\end{aligned}$$

where the set $\mathbf{y}_\ell/\overline{\mathbf{y}}^{\mathbf{t}}<\underline{\mathbf{y}}_{\mathbf{t}},\overline{\mathbf{y}}^{\mathbf{s}}<\underline{\mathbf{y}}_{\mathbf{s}}$ consists of all ℓ-tuple of $[n-1]^{\ell,\neq}$ such that, for $\mathbf{y}=(y_1,\dots,y_\ell)$, $\max_i y_i-\min_i y_i$ is less than or equal to m_n-1 and larger than ℓ. Assume now that for any $k,q\leqslant m_n$, $q>k$ all elements in $(\mathbf{y}_\ell,\mathbf{z}_{r-\ell},\mathbf{t}_{k-r},\mathbf{s}_{q-r})$ are less than or equal to m_n. Then,

$$\begin{aligned}
u_k \star_r^\ell u_q(\mathbf{z}_{r-\ell},\mathbf{t}_{k-r},\mathbf{s}_{q-r}) &= \frac{1}{k!q!}\sum_{\mathbf{y}_\ell\in[m_n]^{\ell,\neq}}\underline{\mathbf{y}}_{\mathbf{t}}\underline{\mathbf{y}}_{\mathbf{s}}(1-p_n)^2p_n^{2m_n-k-q+\ell}+\mathrm{o}(p_n^{2m_n-q+\ell})\\
&\leqslant\frac{1}{k!q!}\binom{m_n}{\ell}^2(1-p_n)^2p_n^{2m_n-k-q+\ell}+\mathrm{o}(p_n^{2m_n-q+\ell}).
\end{aligned}$$

We can see that the closer k and q are to m_n (whatever $\ell\leqslant k$), the smaller $u_k\star_r^\ell u_q$. Moreover, if ℓ is close to zero, $u_k\star_r^\ell u_q$ is even smaller. Then using (B.2) together with the previous remark, we have

$$\begin{aligned}
\sqrt{\mathbf{E}\left[\left|\mathrm{var}[\mathrm{F}]-\int_{\mathbb{N}}\widetilde{\mathrm{D}}_t\mathrm{F}(-\mathrm{D}_t\mathrm{L}^{-1}\mathrm{F})\mathrm{d}\nu(t)\right|^2\right]} &\leqslant \mathfrak{c}\big(\mathfrak{m}_1^\star+\mathfrak{m}_2^\star\big)\\
&\leqslant \mathfrak{c}u_1\star_1^1 u_1\leqslant m_n^2(1-p_n)^2p_n^{2m_n-1}+\mathrm{o}(p_n^{2m_n}).
\end{aligned}$$

where $\mathfrak{c}$ is a generic constant that may change from one line to another. Hence the result. □

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. MR0972770

[2] R. Arratia, L. Goldstein, and L. Gordon. Poisson approximation and the Chen-Stein method. *Statistical Science*, pages 403–424, 1990. MR1092983

[3] S. Attal, J. Deschamps, and C. Pellegrini. Complex obtuse random walks and their continuous-time limits. *Probability Theory and Related Fields*, 165(1):65–116, 2016. MR3500268

[4] S. Attal and M. Émery. Équations de structure pour des martingales vectorielles. In *Séminaire de Probabilités XXVIII*, pages 256–278. Springer, 1994. MR1329117

[5] A. Barbour and L. H. Chen. *An introduction to Stein's method*, volume 4 of *Lecture Notes Series*. National University of Singapore, 2005. MR2235447

[6] A. Barbour, L. H. Chen, and W.-L. Loh. Compound Poisson approximation for nonnegative random variables via Stein's method. *The Annals of Probability*, pages 1843–1866, 1992. MR1188044

[7] A. Barbour and O. Chryssaphinou. Compound Poisson approximation: a user's guide. *Annals of Applied Probability*, pages 964–1002, 2001. MR1865030

[8] A. Barbour, L. Holst, and S. Janson. *Poisson approximation*. The Clarendon Press Oxford Univ., 1992. MR1163825

[9] A. Barbour and S. Utev. Solving the Stein equation in compound Poisson approximation. *Advances in Applied Probability*, pages 449–475, 1998. MR1642848

[10] K. Bichteler. Malliavin calculus for processes with jumps. *Stochastics Monographs*, 1987. MR1008471

[11] S. Bobkov, F. Götze, and H. Sambale. Higher order concentration of measure. *Communications in Contemporary Mathematics*, 21(03):1850043, 2019. MR3947067

[12] L. H. Chen. On the convergence of Poisson binomial to Poisson distr. *Ann. of Proba.*, 2(1):178–180, 1974. MR0370693

[13] H. Chernoff. A note on an inequality involving the Normal distribution. *Ann. of Prob.*, 533–535,1981. MR0614640

[14] R. Dalang, A. Morton, and W. Willinger. Equivalent martingale measures and no-arbitrage in stochastic securities market models. *Stochastics: An Int. Journal of Prob. and Sto. Pr.*, 29(2):185–201, 1990. MR1041035

[15] L. Decreusefond and H. Halconruy. Malliavin and Dirichlet structures for independent random variables. *Stochastic Processes and their Applications*, 129(8):2611–2653, Aug 2019. MR3980139

[16] L. Decreusefond, M. Schulte, and C. Thäle. Functional Poisson approximation in Kantorovich–Rubinstein distance with applications to U-statistics and stochastic geometry. *Ann. of Proba.*, 44(3):2147–2197, 2016. MR3502603

[17] C. Döbler, G. Peccati, et al. The fourth moment theorem on the Poisson space. *The Annals of Probability*, 46(4):1878–1916, 2018. MR3813981

[18] M. Duerinckx. On the size of chaos via Glauber calculus in the classical mean-field dynamics. *Communications in Mathematical Physics*, 382(1):613–653, 2021. MR4223483

[19] M. Duerinckx, A. Gloria, and F. Otto. The structure of fluctuations in stochastic homogenization. *Communications in Mathematical Physics*, pages 1–48, 2020. MR4107930

[20] T. Erhardsson. Stein's method for Poisson and compound Poisson. *An introduction to Stein's method*, 4. MR2235449

[21] I. Flint, N. Privault, and G. L. Torrisi. Functional inequalities for marked point processes. *Electronic Journal of Probability*, 24:1–40, 2019. MR4029419

[22] I. Flint, N. Privault, and Giovanni L. Torrisi. Bounds in total variation distance for discrete-time processes on the sequence space. *Potential Analysis*, 52(2):223–243, 2020. MR4064319

[23] H. Föllmer and M. Schweizer. Hedging of contingent claims under incomplete information. *Applied stochastic analysis*, 5(389-414):19–31, 1991. MR1108430

[24] H. Föllmer and D. Sondermann. Hedging of non-redundant contingent claims. 1985. MR0902885

[25] U. Franz and T. Hamdi. Stochastic analysis for obtuse random walks. *Journal of Theoretical Probability*, 28(2):619–649, 2015. MR3370668

[26] 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. MR1363330

[27] H. Halconruy. The insider problem in the trinomial model: a discrete-time jump process approach. *arXiv preprint arXiv:2106.15208*, 2021.

[28] I. Karatzas, D. L. Ocone, and J. L. Jinlu. An extension of Clark's formula. *Stochastics: An International Journal of Probability and Stochastic Processes*, 37(3):127–131, 1991. MR1148344

[29] R. Lachièze-Rey and G. Peccati. Fine Gaussian fluctuations on the Poisson space II: rescaled kernels, marked processes and geometric U-statistics. *Stoch. Proc. and App.*, 123(12):4186–4218, 2013. MR3096352

[30] G. Last. Stochastic analysis for Poisson processes. In *Stoch. anal. for Poisson point proc.* Springer, 2016. MR3585396

[31] G. Last, G. Peccati, and M. Schulte. Normal approximation on Poisson spaces: Mehler's formula, second order Poincaré inequalities and stabilization. August 2016. MR3520016

[32] 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. MR2802467

[33] G. Last and M. Penrose. *Lectures on the Poisson Process*. Institute of Mathematical Statistics Textbooks. Cambridge University Press, 2017. MR3791470

[34] M. Métivier. Semimartingales, volume 2 of de Gruyter Studies in Mathematics, 1982. MR0688144

[35] J. Nash. Continuity of solutions of parabolic and elliptic equations. *American Journal of Mathematics*, 80(4):931–954, 1958. MR0100158

[36] I. Nourdin and G. Peccati. Noncentral convergence of multiple integrals. *The Annals of Probability*, 37(4):1412–1426, 2009. MR2546749

[37] I. Nourdin and G. Peccati. *Normal Approximations with Malliavin Calculus: From Stein's Method to Universality*. Cambridge University Press, 2012. MR2962301

[38] D. Nualart. *The Malliavin Calculus and Related Topics*. Springer-Verlag Berlin Heidelberg, 2 ed., 2006. MR2200233

[39] D. Nualart and W. Schoutens. Chaotic and predictable representations for Lévy processes. *Stochastic Processes and their Applications*, 90(1):109–122, 2000. MR1787127

[40] 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. MR1071538

[41] G. Di Nunno, B. Øksendal, and F. Proske. White noise analysis for Lévy processes. *Journal of Functional Analysis*, 206:109–148, 2004. MR2024348

[42] D. L. Ocone and I. Karatzas. A generalized Clark representation formula, with application to optimal portfolios. *Stochastics and Stochastic Reports*, 34(3-4):187–220, 1991. MR1124835

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

[44] G. Peccati and C. Zheng. Universal Gaussian fluctuations on the discrete Poisson chaos. *Bernoulli*, 20(2):697–715, 2014. MR3178515

[45] N. Privault. Chaotic and variational calculus in discrete and continuous time for the Poisson process. *Stoch. & Stoch. Rep.*, 51:83–109, 1994. MR1380764

[46] N. Privault. *Stochastic Analysis in discrete and Continuous settings: with normal martingales*. Springer, 2009. MR2531026

[47] N. Privault. *Stochastic finance: An introduction with market examples*. CRC Press, 2013. MR3202743

[48] N. Privault and G. Serafin. Stein approximation for functionals of independent random sequences. *Electronic Journal of Probability*, 23:1–34, 2018. MR3761564

[49] 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. MR1842157

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

[51] S. Schbath. Compound Poisson approximation of word counts in DNA sequences. *ESAIM: PS*, 1:1–16, 1997. MR1382515

[52] M. Schweizer. Option hedging for semimartingales. *Stoch. Proc. and their App.*, 37(2):339–363, 1991. MR1102880

[53] M. Schweizer. Mean-variance hedging for general claims. *Ann. of App. Proba.*, pages 171–179, 1992. MR1143398

[54] M. Schweizer. Variance-optimal hedging in discrete time. *Math.of Operations Research*, 20(1):1–32, 1995. MR1320445

[55] C. Stein. A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. In *Pro. of the 6th Berkeley Symp. on Math. Stat. Proba. p. 583–602*. Univ. of Calif. Press, 1972. MR0402873

[56] D. Stroock. Homogeneous chaos revisited. In *Séminaire de Probabilités XXI*, pages 1–7. Springer, 1987. MR0941972

[57] 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. MR3634270

[58] 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. MR1800540

Acknowledgments. I am also very grateful to Giovanni Peccati for motivating discussions and for helpful advice on writing, and to Antonin Bourgeois, Valentin Garino and Pierre Perruchaud for their help with language issues.