

The Malliavin-Stein method for normal random walks with dependent increments

Ian Flint ^{*} Nicolas Privault [†] Giovanni Luca Torrisi [‡]

Abstract

We derive bounds on the Kolmogorov distance between the distribution of a random functional of a $\{0, 1\}$ -valued random sequence and the normal distribution. Our approach, which relies on the general framework of stochastic analysis for discrete-time normal martingales, extends existing results obtained for independent Bernoulli (or Rademacher) sequences. In particular, we obtain Kolmogorov distance bounds for the sum of normalized random sequences without any independence assumption.

Key words: Normal approximation; central limit theorem; Stein-Chen method; Malliavin-Stein method; Berry-Esseen bound; Kolmogorov distance; normal random walks.

Mathematics Subject Classification: 60F05, 60H07, 60G50, 05C80.

1 Introduction

The Malliavin-Stein method has been introduced in [10] to derive bounds on the distances between probability laws for the normal approximation of functionals of Gaussian random fields, and extended to functionals of Poisson random measures in [12], [13]. Functionals of discrete-time independently distributed Rademacher sequences have been treated using the Wasserstein distance in e.g. [11] for functionals of symmetric sequences, and in [15] in the case of not necessarily symmetric sequences. Bounds in the Kolmogorov distance have been obtained in [8], [9], and bounds in the total variation distance have been derived for the Poisson approximation in [15] and [7], see also [6] for the Poisson approximation of marked binomial processes.

In this paper, we extend the Kolmogorov bounds of [8], [9] from independent Rademacher sequences to the functionals of a suitable discrete-time normal martingale, see [2]. This allows us to consider functionals of arbitrary, not necessarily independent, binary random sequences $\{Y_n\}_{n \in \mathbb{N}}$ generating a filtration $\{\mathcal{F}_n\}_{n \geq -1}$, $\mathcal{F}_{-1} := \{\emptyset, \Omega\}$, and satisfying the normalization condition

$$\mathbb{E}[Y_n | \mathcal{F}_{n-1}] = 0 \quad \text{and} \quad \text{Var}[Y_n | \mathcal{F}_{n-1}] := \mathbb{E}[Y_n^2 | \mathcal{F}_{n-1}] = 1, \quad n \in \mathbb{N} := \{0, 1, \dots\}$$

^{*}School of BioSciences, the University of Melbourne, Parkville VIC 3010, Australia. e-mail: ian.flint@unimelb.edu.au

[†]Division of Mathematical Sciences, Nanyang Technological University, SPMS-MAS-05-43, 21 Nanyang Link Singapore 637371. e-mail: nprivault@ntu.edu.sg

[‡]Consiglio Nazionale delle Ricerche, Via dei Taurini 19, 00185 Roma, Italy. e-mail: giovanniluca.torrisi@cnr.it

see Proposition 2.1. As an example, the binary sequence $\{Y_n\}_{n \in \mathbb{N}}$ can be driven by a time-inhomogeneous two-state Markov chain, see Example 2.3.

In particular, in Theorem 3.1 and Corollary 3.2 we provide bounds on the Kolmogorov distance d_K between the distribution of a random functional of a binary random sequence and the normal distribution. This construction has also been used to derive bounds in total variation between the distributions of random sequences in [4].

Our approach relies on the construction of discrete multiple stochastic integrals and of Malliavin operators for discrete-time binary normal martingales as in [5], which extends the framework of Chapter 1 in [14] to possibly dependent normalized increments. Those operators are used to formulate covariance identities based on the number (or Ornstein-Uhlenbeck) operator L acting on multiple Wiener-Poisson stochastic integrals and its inverse L^{-1} .

Under such covariance identities, a general estimate is derived in Theorem 3.1 and yields an upper bound on the Kolmogorov distance $d_K(F, \mathcal{N})$ between the distribution of a general functional F of $\{Y_n\}_{n \in \mathbb{N}}$ and the normal distribution. This result is then specialized in Corollary 3.2 as a more explicit Kolmogorov upper bound under additional integrability assumptions.

When F is a first order discrete stochastic integral, the general bound of Theorem 3.1 is considerably simplified in Section 4. As an application, this yields estimates on the Kolmogorov distance between the distribution of a linear functional

$$F_N := \sum_{n=0}^N f_N(n) Y_n$$

of the binary sequence $\{Y_n\}_{n \in \mathbb{N}}$ and the normal distribution, see Corollary 4.1.

The paper is organized as follows. In Section 2, we give some preliminary results on normal random walks. In Section 3 we provide our general Kolmogorov upper bound, and we make it more explicit in some specific settings. The above-mentioned applications to discrete-time Markov chains with finite state space and random walks are given in Section 4.

2 Stochastic analysis of normal random walks

In this section we introduce some elements of stochastic analysis for normal random walks, and refer the reader to [2] and [5] for more insight into this subject.

2.1 Normal random walks

Consider the sequence space $\Omega := \{0, 1\}^{\mathbb{N}}$ with its canonical $\{0, 1\}$ -valued coordinate maps $\pi_n : \Omega \rightarrow \{0, 1\}$ defined by $\pi_n((\omega_k)_{k \in \mathbb{N}}) := \omega_n$, for $n \in \mathbb{N}$. We endow Ω with the filtration $\{\mathcal{F}_n\}_{n \in \{-1\} \cup \mathbb{N}}$ defined by

$$\mathcal{F}_{-1} := \{\emptyset, \Omega\}, \quad \mathcal{F}_n := \sigma\{\pi_0, \dots, \pi_n\}, \quad n \in \mathbb{N}, \quad (2.1)$$

and let \mathbb{P} be any probability measure on (Ω, \mathcal{F}) such that

$$0 < p_n := \mathbb{P}(\pi_n = 1 \mid \mathcal{F}_{n-1}) < 1, \quad n \in \mathbb{N}, \quad \mathbb{P}\text{-almost surely.} \quad (2.2)$$

In what follows, we consider the predictable processes $\{v_n^{(0)}\}_{n \in \mathbb{N}}$ and $\{v_n^{(1)}\}_{n \in \mathbb{N}}$ defined by

$$v_n^{(0)} := -\sqrt{\frac{q_n}{p_n}}, \quad v_n^{(1)} := \sqrt{\frac{p_n}{q_n}}, \quad n \in \mathbb{N}, \quad (2.3)$$

where $q_n := 1 - p_n$, $n \in \mathbb{N}$.

Proposition 2.1. *The process $\{Y_0 + \dots + Y_n\}_{n \in \mathbb{N}}$ defined by its normalized increments*

$$Y_n(\omega) := v_n^{(\pi_n(\omega))}(\omega), \quad n \in \mathbb{N}, \quad \omega \in \Omega,$$

is a one-dimensional normal martingale, i.e.

$$\mathbb{E}[Y_n \mid \mathcal{F}_{n-1}] = 0 \quad \text{and} \quad \text{Var}[Y_n \mid \mathcal{F}_{n-1}] := \mathbb{E}[Y_n^2 \mid \mathcal{F}_{n-1}] = 1, \quad n \in \mathbb{N}. \quad (2.4)$$

Proof. From (2.3) we have the relations

$$|v_n^{(0)}|^2 = \frac{1 - p_n}{p_n}, \quad |v_n^{(1)}|^2 = \frac{1 - q_n}{q_n}, \quad n \in \mathbb{N}, \quad (2.5)$$

which yield the structure equations

$$p_n v_n^{(0)} + q_n v_n^{(1)} = 0, \quad p_n |v_n^{(0)}|^2 + q_n |v_n^{(1)}|^2 = 1, \quad n \in \mathbb{N}, \quad (2.6)$$

and (2.4). □

Lemma 2.1. *Let $a^{(i)}, b^{(i)} : \Omega \rightarrow \mathbb{R}$, $i \in \{0, 1\}$, be two random variables. For any $n \in \mathbb{N}$, we have*

$$(p_n a^{(0)} v_n^{(0)} + q_n a^{(1)} v_n^{(1)})(p_n b^{(0)} v_n^{(0)} + q_n b^{(1)} v_n^{(1)}) = p_n q_n (a^{(0)} - a^{(1)})(b^{(0)} - b^{(1)}),$$

\mathbb{P} -almost surely.

Proof. By the definition of the processes $\{v_n^{(0)}\}_{n \in \mathbb{N}}$ and $\{v_n^{(1)}\}_{n \in \mathbb{N}}$, and (2.5), we have

$$\begin{aligned} & (p_n a^{(0)} v_n^{(0)} + q_n a^{(1)} v_n^{(1)})(p_n b^{(0)} v_n^{(0)} + q_n b^{(1)} v_n^{(1)}) \\ &= p_n^2 a^{(0)} b^{(0)} |v_n^{(0)}|^2 + q_n^2 a^{(1)} b^{(1)} |v_n^{(1)}|^2 + v_n^{(0)} v_n^{(1)} p_n q_n (a^{(0)} b^{(1)} + a^{(1)} b^{(0)}) \\ &= p_n q_n (a^{(0)} - a^{(1)})(b^{(0)} - b^{(1)}), \quad n \in \mathbb{N}. \end{aligned}$$

□

We also note that

$$\{\omega \in \Omega : \pi_n(\omega) = i\} = \{\omega \in \Omega : Y_n(\omega) = v_n^{(i)}(\omega)\}, \quad i = 0, 1, \quad n \in \mathbb{N},$$

hence $\mathcal{F}_n = \sigma\{Y_0, \dots, Y_n\}$, $n \in \mathbb{N}$, and

$$\mathbb{P}(Y_n = v_n^{(0)} \mid \mathcal{F}_{n-1}) = p_n = \mathbb{P}(\pi_n = 0 \mid \mathcal{F}_{n-1}), \quad \mathbb{P}(Y_n = v_n^{(1)} \mid \mathcal{F}_{n-1}) = q_n = \mathbb{P}(\pi_n = 1 \mid \mathcal{F}_{n-1}),$$

$n \in \mathbb{N}$.

Remark 2.2. *The above construction admits an extension to so-called obtuse systems in \mathbb{R}^d , see [2], [1], in which case $(Y_0 + \dots + Y_n)_{n \in \{0, \dots, N\}}$ is a d -dimensional normal martingale, i.e.*

$$\mathbb{E}[Y_n \mid \mathcal{F}_{n-1}] = 0 \text{ and } \text{Var}[Y_n \mid \mathcal{F}_{n-1}] := \mathbb{E}[Y_n \otimes Y_n \mid \mathcal{F}_{n-1}] = \mathbf{I}_n, \quad n \in \{0, \dots, N\}.$$

Next, we note that our construction allows the binary sequence $\{Y_n\}_{n \in \mathbb{N}}$ to be driven by a time-inhomogeneous two-state Markov chain.

Example 2.3. *When*

$$\mathbb{P}(\pi_n = 1 \mid \mathcal{F}_{n-1}) = \mathbb{P}(\pi_n = 1 \mid \pi_{n-1}), \quad n \in \mathbb{N}, \quad \mathbb{P}\text{-almost surely},$$

the process $\{\pi_n\}_{n \in \mathbb{N}}$ becomes a $\{0, 1\}$ -valued time-inhomogeneous Markov chain with transition probabilities

$$P_{i,j}^{(n)} := \mathbb{P}(\pi_n = j \mid \pi_{n-1} = i), \quad i, j = 0, 1,$$

and initial distribution $p_0 = \mathbb{P}(\pi_0 = 0)$, $q_0 = \mathbb{P}(\pi_0 = 1)$, and we have

$$\{Y_n\}_{n \in \mathbb{N}} := \left\{ -\mathbf{1}_{\{\pi_n=0\}} \sqrt{\frac{\mathbb{P}(\pi_n = 1 \mid \pi_{n-1})}{\mathbb{P}(\pi_n = 0 \mid \pi_{n-1})}} + \mathbf{1}_{\{\pi_n=1\}} \sqrt{\frac{\mathbb{P}(\pi_n = 0 \mid \pi_{n-1})}{\mathbb{P}(\pi_n = 1 \mid \pi_{n-1})}} \right\}_{n \in \mathbb{N}}.$$

The next example considers a binomial asset price model with dependence.

Example 2.4. Given a real-valued sequence $(\sigma(n))_{0 \leq n \leq N}$, $\sigma(0) := 0$, we consider a binomial asset pricing model. For $S_0 > 0$ an initial positive constant, the discrete-time risky asset price process $(S_N)_{N \in \mathbb{N}}$ is defined recursively by the price dynamics

$$S_n = S_{n-1} e^{\sigma(n) Y_n / \sqrt{N+1} - \sigma^2(n)/(2(N+1))} = \begin{cases} S_{n-1} e^{\sigma(n) \sqrt{p_n / ((N+1)q_n)} - \sigma^2(n)/(2(N+1))}, & \pi_n = 1, \\ S_{n-1} e^{-\sigma(n) \sqrt{q_n / ((N+1)p_n)} - \sigma^2(n)/(2(N+1))}, & \pi_n = 0, \end{cases}$$

$n = 1, \dots, N$, yielding

$$S_N = S_0 \exp \left(\frac{1}{\sqrt{N+1}} \sum_{n=0}^N \sigma(n) Y_n - \frac{1}{2(N+1)} \sum_{n=0}^N \sigma^2(n) \right).$$

2.2 Discrete multiple stochastic integrals

We denote by κ the counting measure on \mathbb{N} , set $\ell^2(\mathbb{N}^n) := L^2(\mathbb{N}^n, \mathcal{P}(\mathbb{N})^{\otimes n}, \kappa^{\otimes n})$ for $n \in \mathbb{N}^* := \mathbb{N} \setminus \{0\}$, with $\ell^2(\mathbb{N})^{\otimes 0} := \mathbb{R}$, and refer to the elements of $\ell^2(\mathbb{N}^n)$ as kernels. By $\ell^2(\mathbb{N})^{\text{on}}$ we denote the class of symmetric kernels and by $\ell_0^2(\mathbb{N})^{\text{on}}$ the family of symmetric kernels which vanish on the diagonals, i.e., which vanish on the complement of the set

$$\Delta_n := \{(i_1, \dots, i_n) \in \mathbb{N}^n : i_h \neq i_k, h \neq k\}.$$

Let $f_n \in \ell_0^2(\mathbb{N})^{\text{on}}$ be a symmetric kernel vanishing on the diagonal. For $n \in \mathbb{N}^*$, the discrete multiple stochastic integral of order n of f_n is defined by

$$\begin{aligned} J_n(f_n) &:= \sum_{(i_1, \dots, i_n) \in \Delta_n} f_n(i_1, \dots, i_n) Y_{i_1} \cdots Y_{i_n} \\ &= n! \sum_{0 \leq i_1 < \dots < i_n < \infty} f_n(i_1, \dots, i_n) Y_{i_1} \cdots Y_{i_n}, \end{aligned} \quad (2.7)$$

where the random variables Y_n are defined in Proposition 2.1. We also set $J_0(c) := c$ for any $c \in \mathbb{R}$. We call the space spanned by the random variables $J_n(f)$, with $f \in \ell_0^2(\mathbb{N})^{\text{on}}$, the n -th chaos of the normal random walk. Discrete multiple stochastic integrals of different orders are mutually orthogonal and satisfy the isometry relation

$$\mathbb{E}[J_n(f_n) J_m(g_m)] = \mathbf{1}_{\{n=m\}} n! \langle f_n, g_n \rangle_{\ell^2(\mathbb{N})^{\otimes n}}, \quad (2.8)$$

for any couple of symmetric kernels f_n, g_m , $m, n \in \mathbb{N}^*$ (see Proposition 3.4 in [5]). Discrete multiple stochastic integrals are centered random variables, i.e., $\mathbb{E}[J_n(f_n)] = 0$, for any symmetric kernel f_n , $n \in \mathbb{N}^*$ (see the proof of Proposition 5.1 in [5]). Finally, we recall that, for any real-valued random variable $F \in L^2(\Omega, \mathcal{F}, \mathbb{P})$, where $\mathcal{F} := \bigvee_{n \in \mathbb{N}} \mathcal{F}_n$, the chaotic decomposition

$$F = \mathbb{E}[F] + \sum_{n=1}^{\infty} J_n(f_n) \quad (2.9)$$

holds, for uniquely determined symmetric kernels f_n (see Theorem 5.7 in [5]).

2.3 Malliavin operators

We define the discrete gradient of a random variable $F : \Omega \rightarrow \mathbb{R}$ as

$$D_n F := p_n v_n^{(0)} F_n^0 + q_n v_n^{(1)} F_n^1 = \sqrt{p_n q_n} (F_n^1 - F_n^0), \quad n \in \mathbb{N}, \quad (2.10)$$

where $F_n^i(\omega) := F(\omega_n^i)$ and $\omega_n^i := (\omega_0, \dots, \omega_{n-1}, i, \omega_{n+1}, \dots)$, $i = 0, 1$, $n \in \mathbb{N}$. The following proposition holds (see [5]).

Proposition 2.5. *We have:*

(i) *For random variables $F, G : \Omega \rightarrow \mathbb{R}$,*

$$D_n(FG) = F D_n G + G D_n F - \sqrt{p_n q_n} [(F - F_n^0)(G - G_n^0) - (F - F_n^1)(G - G_n^1)], \quad n \in \mathbb{N}.$$

(ii) *For any $n \in \mathbb{N}^*$ and symmetric kernel f_n ,*

$$D_k J_n(f_n) = n J_{n-1}(f_n(*, k)), \quad k \in \mathbb{N}.$$

In particular, we have the following corollary.

Corollary 2.6. *For random variables $F, G : \Omega \rightarrow \mathbb{R}$,*

$$D_n(FG) = F D_n G + G D_n F - \frac{2\pi_n - 1}{\sqrt{p_n q_n}} (D_n F)(D_n G).$$

Proof. By (2.10) and Proposition 2.5(i) we have

$$\begin{aligned} D_n(FG) &= F D_n G + G D_n F - \sqrt{p_n q_n} [(F_n^1 - F_n^0)(G_n^1 - G_n^0) \mathbf{1}_{\{\pi_n=1\}} - (F_n^1 - F_n^0)(G_n^1 - G_n^0) \mathbf{1}_{\{\pi_n=0\}}] \\ &= F D_n G + G D_n F - \frac{1}{\sqrt{p_n q_n}} [(D_n F)(D_n G) \mathbf{1}_{\{\pi_n=1\}} - (D_n F)(D_n G) \mathbf{1}_{\{\pi_n=0\}}] \\ &= F D_n G + G D_n F - \frac{2\pi_n - 1}{\sqrt{p_n q_n}} (D_n F)(D_n G). \end{aligned}$$

□

The L^2 -domain of D , denoted by $\text{Dom}(D)$, is defined by

$$\begin{aligned} \text{Dom}(D) &:= \left\{ F \in L^2(\Omega, \mathcal{F}, \mathbb{P}) : \mathbb{E} \left[\sum_{n \in \mathbb{N}} |D_n F|^2 \right] < \infty \right\} \\ &= \left\{ F = \sum_{n \in \mathbb{N}} J_n(f_n) \in L^2(\Omega, \mathcal{F}, \mathbb{P}) : \sum_{n=1}^{\infty} n(n!) \|f_n\|_{\ell^2(\mathbb{N})^{\otimes n}}^2 < \infty \right\}. \end{aligned}$$

For $F \in \text{Dom}(D)$, $F = \sum_{n \in \mathbb{N}} J_n(f_n)$, we have

$$D_k F = \sum_{n=1}^{\infty} n J_{n-1}(f_n(*, k)). \quad (2.11)$$

Next, we introduce the Ornstein-Uhlenbeck operator L and its (pseudo-)inverse L^{-1} . The L^2 -domain of L , denoted by $\text{Dom}(L)$, is defined by

$$\text{Dom}(L) := \left\{ F = \sum_{n=0}^{\infty} J_n(f_n) \in L^2(\Omega, \mathcal{F}, \mathbb{P}) : \sum_{n=1}^{\infty} n^2 (n!) \|f_n\|_{\ell^2(\mathbb{N})^{\otimes n}}^2 < \infty \right\}.$$

For $F \in \text{Dom}(L)$, $F = \sum_{n=0}^{\infty} J_n(f_n)$, we put

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

For a centered $F = \sum_{n=1}^{\infty} J_n(f_n) \in L^2(\Omega, \mathcal{F}, \mathbb{P})$, we define the (pseudo-)inverse operator of L as

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

Next, we introduce the divergence operator δ . The L^2 -domain of δ , denoted by $\text{Dom}(\delta)$, is defined as follows. Let $u := (u_k)_{k \in \mathbb{N}} \in (L^2(\Omega, \mathcal{F}, \mathbb{P}))^{\mathbb{N}}$ be such that there exists a sequence $g_{n+1} \in \ell_0^2(\mathbb{N})^{\otimes n} \otimes \ell^2(\mathbb{N})$, $n \in \mathbb{N}$, such that

$$u_k := \sum_{n \in \mathbb{N}} J_n(g_{n+1}(*, k)). \quad (2.12)$$

We say that $u \in \text{Dom}(\delta)$ if

$$\sum_{n=0}^{\infty} (n+1)! \|\tilde{g}_{n+1} \mathbb{1}_{\Delta_{n+1}}\|_{\ell^2(\mathbb{N})^{\otimes (n+1)}}^2 < \infty, \quad (2.13)$$

and in this case we define

$$\delta(u) := \sum_{n=0}^{\infty} J_{n+1}(\tilde{g}_{n+1} \mathbb{1}_{\Delta_{n+1}}). \quad (2.14)$$

Here, \tilde{f} denotes the canonical symmetrization of f . Note that, for $u \in \text{Dom}(\delta)$, (2.13) can be rewritten as $\mathbb{E}[\delta(u)^2] < \infty$.

The following proposition provides an integration by parts formula (see Proposition 8.2 in [5]).

Proposition 2.7. *The operator δ is the adjoint of D , i.e., for all $F \in \text{Dom}(D)$ and $u \in \text{Dom}(\delta)$,*

$$\mathbb{E}[F\delta(u)] = \mathbb{E}[\langle DF, u \rangle_{\ell^2(\mathbb{N})}]. \quad (2.15)$$

The next result is standard, and states that the operators D , L and δ are related by the identity $-\delta D = L$.

Proposition 2.8. *For any $F \in \text{Dom}(L)$, we have $F \in \text{Dom}(D)$, $DF \in \text{Dom}(\delta)$ and $-\delta(DF) = LF$.*

Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be measurable and $F \in L^2(\Omega, \mathcal{F}, \mathbb{P})$ centered and such that $f(F) \in \text{Dom}(D)$. Then, by Propositions 2.7 and 2.8,

$$\mathbb{E}[Ff(F)] = \mathbb{E}[L(L^{-1}F)f(F)] = -\mathbb{E}[\delta(D(L^{-1}F))f(F)] = \mathbb{E}[\langle Df(F), -D(L^{-1}F) \rangle_{\ell^2(\mathbb{N})}]. \quad (2.16)$$

This relation will be crucial when we develop the Malliavin-Stein method. The next technical lemma is exploited to prove the Skorohod isometry.

Lemma 2.9. *For any $u \in \text{Dom}(\delta)$ and any $k \in \mathbb{N}$, we have $D_k u \in \text{Dom}(\delta)$.*

Proof. Let $u \in \text{Dom}(\delta)$. For $k \in \mathbb{N}$ fixed and any $\ell \in \mathbb{N}$, we have

$$\begin{aligned} D_k u_\ell &= \sum_{n=1}^{\infty} n J_{n-1}(g_{n+1}(*, \ell, k)) = \sum_{n=0}^{\infty} J_n((n+1)g_{n+2}(*, \ell, k)) \\ &= \sum_{n=0}^{\infty} J_n(f_{n+1}(*, \ell)), \end{aligned} \quad (2.17)$$

where $f_{n+1}(*, \ell) := (n+1)g_{n+2}(*, \ell, k)$ (the dependence on k is not made explicit). We have to prove that

$$\sum_{n=0}^{\infty} (n+1)! \|\tilde{f}_{n+1} \mathbf{1}_{\Delta_{n+1}}\|_{\ell^2(\mathbb{N})^{\otimes(n+1)}}^2 < \infty.$$

We have

$$\sum_{n=0}^{\infty} (n+1)! \|\tilde{f}_{n+1} \mathbf{1}_{\Delta_{n+1}}\|_{\ell^2(\mathbb{N})^{\otimes(n+1)}}^2 = \sum_{n=0}^{\infty} (n+1)!(n+1)^2 \|h_{n+1} \mathbf{1}_{\Delta_{n+1}}\|_{\ell^2(\mathbb{N})^{\otimes(n+1)}}^2, \quad (2.18)$$

where, letting \mathcal{S}_{n+1} be the permutation group on $\{1, \dots, n+1\}$,

$$\begin{aligned} h_{n+1}(x_1, \dots, x_{n+1}) \mathbf{1}_{\Delta_{n+1}}(x_1, \dots, x_{n+1}) \\ = \mathbf{1}_{\Delta_{n+1}}(x_1, \dots, x_{n+1}) \frac{1}{(n+1)!} \sum_{\sigma \in \mathcal{S}_{n+1}} g_{n+2}(x_{\sigma(1)}, \dots, x_{\sigma(n)}, x_{\sigma(n+1)}, k). \end{aligned} \quad (2.19)$$

Note that

$$\begin{aligned} &\|h_{n+1} \mathbf{1}_{\Delta_{n+1}}\|_{\ell^2(\mathbb{N})^{\otimes(n+1)}}^2 \\ &= \sum_{(x_1, \dots, x_{n+1}) \in \mathbb{N}^{n+1}} |h_{n+1}(x_1, \dots, x_{n+1})|^2 \mathbf{1}_{\Delta_{n+1}}(x_1, \dots, x_{n+1}) \\ &= \frac{1}{[(n+1)!]^2} \sum_{(x_1, \dots, x_{n+1}) \in \mathbb{N}^{n+1}} \mathbf{1}_{\Delta_{n+1}}(x_1, \dots, x_{n+1}) \left| \sum_{\sigma \in \mathcal{S}_{n+1}} g_{n+2}(x_{\sigma(1)}, \dots, x_{\sigma(n)}, x_{\sigma(n+1)}, k) \right|^2. \end{aligned} \quad (2.20)$$

Using the fact that g_{n+2} is symmetric with respect to the first n variables, for a fixed $(x_1, \dots, x_{n+1}) \in \mathbb{N}^{n+1}$, we have

$$\sum_{\sigma \in \mathcal{S}_{n+1}} g_{n+2}(x_{\sigma(1)}, \dots, x_{\sigma(n)}, x_{\sigma(n+1)}, k) = n! \sum_{i=1}^{n+1} g_{n+2}(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_{n+1}, x_i, k).$$

Inserting this expression into (2.20) we have

$$\begin{aligned} \|h_{n+1} \mathbf{1}_{\Delta_{n+1}}\|_{\ell^2(\mathbb{N})^{\otimes(n+1)}}^2 &= \frac{1}{(n+1)^2} \sum_{(x_1, \dots, x_{n+1}) \in \mathbb{N}^{n+1}} \mathbf{1}_{\Delta_{n+1}}(x_1, \dots, x_{n+1}) \\ &\quad \times \left| \sum_{i=1}^{n+1} g_{n+2}(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_{n+1}, x_i, k) \right|^2. \end{aligned}$$

By this relation and (2.18), we have

$$\begin{aligned} \sum_{n=0}^{\infty} (n+1)! \|\tilde{f}_{n+1} \mathbf{1}_{\Delta_{n+1}}\|_{\ell^2(\mathbb{N})^{\otimes(n+1)}}^2 &= \sum_{n=0}^{\infty} (n+1)! \sum_{(x_1, \dots, x_{n+1}) \in \mathbb{N}^{n+1}} \mathbf{1}_{\Delta_{n+1}}(x_1, \dots, x_{n+1}) \\ &\quad \times \left| \sum_{i=1}^{n+1} g_{n+2}(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_{n+1}, x_i, k) \right|^2, \end{aligned}$$

and this is a finite quantity due to (2.13). □

The next proposition provides the Skorohod formula for δ . Note that this formula becomes indeed an isometry if u_k is \mathcal{F}_{k-1} -measurable.

Proposition 2.10. *For all $u \in \text{Dom}(\delta)$, we have*

$$\mathbb{E}[\delta(u)^2] = \mathbb{E}[\|u\|_{\ell^2(\mathbb{N})}^2] + \mathbb{E}\left[\sum_{k,l \geq 0} D_l u_k D_k u_l\right] - 2\mathbb{E}\left[\sum_{k \geq 0} (D_k u_k)^2\right].$$

In particular, if u_k is \mathcal{F}_{k-1} -measurable for any k , then

$$\mathbb{E}[\delta(u)^2] = \mathbb{E}[\|u\|_{\ell^2(\mathbb{N})}^2],$$

indeed in such a case $D_l u_k D_k u_l = 0$ for any k, l .

Proof. By (2.14), (2.7) and (2.11), for any $u \in \text{Dom}(\delta)$ we have

$$\begin{aligned} \delta(u) &= \sum_{n=0}^{\infty} J_{n+1}(\tilde{g}_{n+1} \mathbf{1}_{\Delta_{n+1}}) \\ &= \sum_{n=0}^{\infty} \sum_{(i_1, \dots, i_{n+1}) \in \Delta_{n+1}} \tilde{g}_{n+1}(i_1, \dots, i_{n+1}) Y_{i_1} \cdots Y_{i_{n+1}} \end{aligned}$$

$$\begin{aligned}
&= \sum_{k \geq 0} \left(\sum_{n=0}^{\infty} \sum_{(i_1, \dots, i_n) \in \Delta_n} \tilde{g}_{n+1}(i_1, \dots, i_n, k) Y_{i_1} \cdots Y_{i_n} \cdot Y_k \right. \\
&\quad \left. - \sum_{n=1}^{\infty} n \sum_{(i_1, \dots, i_{n-1}) \in \Delta_{n-1}} \tilde{g}_{n+1}(i_1, \dots, i_{n-1}, k, k) Y_{i_1} \cdots Y_{i_{n-1}} \cdot |Y_k|^2 \right) \\
&= \sum_{k \geq 0} (u_k Y_k - (D_k u_k) |Y_k|^2). \tag{2.21}
\end{aligned}$$

By Lemma 2.9 $D_k u \in \text{Dom}(\delta)$ and so from the above formula, we deduce, in particular, that

$$\delta(D_k u) = \sum_{l \geq 0} [Y_l D_k u_l - (D_l D_k u_l) |Y_l|^2]. \tag{2.22}$$

Note that $Y_l = J_1(\mathbf{1}_{\{l\}})$ and so $D_k Y_l = \mathbf{1}_{\{k=l\}}$. Therefore, by Corollary 2.6 and (2.22), we have

$$\begin{aligned}
D_k \delta(u) &= D_k \left(\sum_{l \geq 0} [u_l Y_l - (D_l u_l) |Y_l|^2] \right) \\
&= \sum_{l \geq 0} \left[(D_k u_l) Y_l + u_l D_k Y_l - \frac{2\pi_k - 1}{\sqrt{p_k q_k}} (D_k u_l) (D_k Y_l) - (D_k D_l u_l) |Y_l|^2 \right. \\
&\quad \left. - (D_l u_l) D_k |Y_l|^2 + \frac{2\pi_k - 1}{\sqrt{p_k^{(0)} p_k^{(1)}}} (D_k D_l u_l) (D_k |Y_l|^2) \right] \\
&= \sum_{l \geq 0} \left[(D_k u_l) Y_l - (D_k D_l u_l) |Y_l|^2 + u_l D_k Y_l - \frac{2\pi_k - 1}{\sqrt{p_k q_k}} (D_k u_l) (D_k Y_l) \right. \\
&\quad \left. - (D_l u_l) \left[2Y_l D_k Y_l - \frac{2\pi_k - 1}{\sqrt{p_k q_k}} (D_k Y_l)^2 \right] \right. \\
&\quad \left. + \frac{2\pi_k - 1}{\sqrt{p_k q_k}} (D_k D_l u_l) \left[2Y_l D_k Y_l - \frac{2\pi_k - 1}{\sqrt{p_k q_k}} (D_k Y_l)^2 \right] \right] \\
&= \delta(D_k u) + u_k - \frac{2\pi_k - 1}{\sqrt{p_k q_k}} D_k u_k - (D_k u_k) \left[2Y_k - \frac{2\pi_k - 1}{\sqrt{p_k q_k}} \right] \\
&\quad + \frac{2\pi_k - 1}{\sqrt{p_k q_k}} (D_k D_k u_k) \left[2Y_k - \frac{2\pi_k - 1}{\sqrt{p_k q_k}} \right] \\
&= \delta(D_k u) + u_k - 2(D_k u_k) Y_k,
\end{aligned}$$

where we used that $D_k D_k u_k = 0$. Again by (2.21) and this latter relation, we have

$$\delta(\mathbf{1}_{\{k\}} D_k u) = Y_k D_k u_k - (D_k D_k u_k) |Y_k|^2 = Y_k D_k u_k,$$

and therefore by applying Proposition 2.7 three times,

$$\begin{aligned}
\mathbb{E}[\delta(u)^2] &= \mathbb{E} \left[\sum_{k \geq 0} D_k \delta(u) u_k \right] \\
&= \mathbb{E} \left[\sum_{k \geq 0} (\delta(D_k u) + u_k - 2(D_k u_k) Y_k) u_k \right]
\end{aligned}$$

$$\begin{aligned}
&= \mathbb{E}[\|u\|_{\ell^2(\mathbb{N})}] + \mathbb{E}\left[\sum_{k \geq 0} u_k \delta(D_k u) - 2\delta(\mathbf{1}_{\{k\}} D_k u_k) u_k\right] \\
&= \mathbb{E}[\|u\|_{\ell^2(\mathbb{N})}] + \mathbb{E}\left[\sum_{k, l \geq 0} [D_l u_k D_k u_l - 2\mathbf{1}_{\{k\}} D_k u_k D_l u_k]\right] \\
&= \mathbb{E}[\|u\|_{\ell^2(\mathbb{N})}] + \mathbb{E}\left[\sum_{k, l \geq 0} D_l u_k D_k u_l\right] - 2\mathbb{E}\left[\sum_{k \geq 0} (D_k u_k)^2\right].
\end{aligned}$$

□

Now, we give a generalisation of the integration by parts formula (2.15) which can be applied to square integrable functionals that do not necessarily belong to $\text{Dom}(D)$. This extension of the integration by parts formula plays a role in the Gaussian approximation, see the proof of Corollary 3.2.

Proposition 2.11. *Let $F \in L^2(\Omega, \mathcal{F}, \mathbb{P})$ and let*

$$u = (u_k)_{k \in \mathbb{N}} \in (L^2(\Omega, \mathcal{F}, \mathbb{P}))^{\mathbb{N}}$$

with u_k defined by (2.12) and

$$\sum_{n=0}^{\infty} (n+1)! \|g_{n+1}\|_{\ell^2(\mathbb{N})^{\otimes(n+1)}}^2 < \infty. \quad (2.23)$$

If, for any $k \in \mathbb{N}$, $(D_k F)u_k \geq 0$ \mathbb{P} -almost surely, then $u \in \text{Dom}(\delta)$ and (2.15) holds.

Proof. Although the proof is similar to the proof of Proposition 2.2 in [9], since we are working in a different context, we provide the details. We start by noticing that (2.23) implies (2.13) and therefore $u \in \text{Dom}(\delta)$. Since $F \in L^2(\Omega, \mathcal{F}, \mathbb{P})$, by the chaotic decomposition (2.9) we have

$$F = \mathbb{E}[F] + \sum_{n=1}^{\infty} J_n(f_n)$$

for uniquely determined symmetric kernels f_n . The isometry formula for multiple stochastic integrals (2.8) yields, noticing that $\mathbb{E}[\delta(u)] = 0$,

$$\begin{aligned}
\mathbb{E}[F\delta(u)] &= \mathbb{E}\left[\left(\sum_{n=1}^{\infty} J_n(f_n)\right) \left(\sum_{n=0}^{\infty} J_{n+1}(\tilde{g}_{n+1} \mathbf{1}_{\Delta_{n+1}})\right)\right] \\
&= \mathbb{E}\left[\left(\sum_{n=0}^{\infty} J_{n+1}(f_{n+1})\right) \left(\sum_{n=0}^{\infty} J_{n+1}(\tilde{g}_{n+1} \mathbf{1}_{\Delta_{n+1}})\right)\right] \\
&= \sum_{n \in \mathbb{N}} (n+1)! \langle f_{n+1}, \tilde{g}_{n+1} \mathbf{1}_{\Delta_{n+1}} \rangle_{\ell^2(\mathbb{N})^{\otimes(n+1)}} \\
&= \sum_{n \in \mathbb{N}} (n+1)! \langle f_{n+1}, g_{n+1} \rangle_{\ell^2(\mathbb{N})^{\otimes(n+1)}}.
\end{aligned}$$

On the other hand, again by (2.8) we have

$$\begin{aligned}
\mathbb{E}[\langle DF, u \rangle_{\ell^2(\mathbb{N})}] &= \sum_{k \geq 0} \mathbb{E}[D_k F u_k] \\
&= \sum_{k \geq 0} \mathbb{E} \left[\left(\sum_{n \in \mathbb{N}} (n+1) J_n(f_{n+1}(*, k)) \right) \left(\sum_{n \in \mathbb{N}} J_n(g_{n+1}(*, k)) \right) \right] \\
&= \sum_{k \geq 0} \sum_{n \in \mathbb{N}} (n+1)! \langle f_{n+1}(*, k), g_{n+1}(*, k) \rangle_{\ell^2(\mathbb{N})^{\otimes n}} \\
&= \sum_{n \in \mathbb{N}} (n+1)! \sum_{k \geq 0} \langle f_{n+1}(*, k), g_{n+1}(*, k) \rangle_{\ell^2(\mathbb{N})^{\otimes n}} \tag{2.24} \\
&= \sum_{n \in \mathbb{N}} (n+1)! \langle f_{n+1}, g_{n+1} \rangle_{\ell^2(\mathbb{N})^{\otimes (n+1)}}.
\end{aligned}$$

To complete the proof it remains to justify the exchange between the infinite sums in (2.24). This is guaranteed by Fubini's theorem. Indeed, by a repeated application of the Cauchy-Schwarz inequality and assumption (2.23), one has

$$\sum_{n \in \mathbb{N}} \sum_{k \in \mathbb{N}} |(n+1)! \langle f_{n+1}(*, k), g_{n+1}(*, k) \rangle_{\ell^2(\mathbb{N})^{\otimes n}}| < \infty,$$

see [9] for details.

□

We also recall the following covariance representation formula (see Proposition 9.1 in [5]).

Proposition 2.12. *For $F, G \in \text{Dom}(D)$, we have*

$$\text{Cov}(F, G) = \mathbb{E} \left[\sum_{n \geq 0} \mathbb{E}[D_n G \mid \mathcal{F}_{n-1}] D_n F \right].$$

3 A Berry-Esseen bound

3.1 Stein's equation

Let Φ be the cumulative distribution function of a normal standard random variable. It is well-known (see, e.g., Lemmas 2.2 and 2.3 in [3]) that the unique bounded solution f_x of the Stein equation

$$f'(w) - wf(w) = \mathbf{1}_{(-\infty, x]}(w) - \Phi(x), \quad w, x \in \mathbb{R}$$

is such that $0 < f_x(w) \leq \sqrt{2\pi}/4$, $|f'_x(w)| \leq 1$ for all $w \in \mathbb{R}$ and

$$|(w+u)f_x(w+u) - (w+v)f_x(w+v)| \leq (|w| + \sqrt{2\pi}/4)(|u| + |v|), \quad \text{for all } u, w, v \in \mathbb{R}. \tag{3.1}$$

If we replace w by a random variable F (defined on Ω) in the Stein equation and we take the expectation, we have $\mathbb{E}[f'_x(F) - Ff_x(F)] = \mathbb{P}(F \leq x) - \Phi(x)$ for any $x \in \mathbb{R}$ and so

$$d_K(F, \mathcal{N}) := \sup_{x \in \mathbb{R}} |\mathbb{P}(F \leq x) - \Phi(x)| = \sup_{x \in \mathbb{R}} |\mathbb{E}[f'_x(F) - Ff_x(F)]|, \quad (3.2)$$

where d_K denotes the Kolmogorov distance and \mathcal{N} is a standard normal random variable with cumulative distribution function Φ .

3.2 General Kolmogorov bound

We now establish a bound on the Kolmogorov distance between a centered functional $F \in \text{Dom}(D)$ and \mathcal{N} .

Theorem 3.1. *Let $F \in \text{Dom}(D)$ be such that $\mathbb{E}[F] = 0$ and let G be a real-valued random variable on Ω such that*

$$\text{Cov}(F, f_x(F)) = \mathbb{E}[\langle DG, Df_x(F) \rangle_{\ell^2(\mathbb{N})}], \quad x \in \mathbb{R}, \quad (3.3)$$

where f_x is the solution of the Stein equation (see Remark 3.4 below for possible choices of G). Then

$$d_K(F, \mathcal{N}) \leq \mathbb{E}[|1 - \langle DG, DF \rangle_{\ell^2(\mathbb{N})}|] + \frac{1}{2} \sum_{n \in \mathbb{N}} \mathbb{E} \left[\left(|F| + \frac{\sqrt{2\pi}}{4} \right) |G_n^1 - G_n^0| |D_n F|^2 \right] \\ + \sup_{x \in \mathbb{R}} \sum_{n \in \mathbb{N}} \mathbb{E} [|G_n^1 - G_n^0| |D_n F D_n \mathbf{1}_{\{F > x\}}|].$$

In particular, choosing $G := -L^{-1}F$ (see Remark 3.4 below), Theorem 3.1 shows that

$$d_K(F, \mathcal{N}) \leq \mathbb{E}[|1 - \langle -DL^{-1}F, DF \rangle_{\ell^2(\mathbb{N})}|] + \frac{1}{2} \sum_{n \in \mathbb{N}} \mathbb{E} \left[(p_n q_n)^{-1/2} \left(|F| + \frac{\sqrt{2\pi}}{4} \right) |D_n L^{-1}F| |D_n F|^2 \right] \\ + \sup_{x \in \mathbb{R}} \sum_{n \in \mathbb{N}} \mathbb{E} \left[(p_n q_n)^{-1/2} |D_n L^{-1}F| |D_n F D_n \mathbf{1}_{\{F > x\}}| \right], \quad (3.4)$$

which extends Theorem 3.1 of [8] and Proposition 4.1 of [9] to possibly non-independent random sequences. The next corollary specializes the above bound under an additional integrability assumption, and similarly extends Theorem 4.1 of [9].

Corollary 3.2. *Let $F \in \text{Dom}(D)$ be such that $\mathbb{E}[F] = 0$, and let*

$$u_k := (p_k q_k)^{-1/2} D_k F |D_k L^{-1}F|, \quad k \in \mathbb{N}, \quad (3.5)$$

be such that $u_k \in L^2(\Omega, \mathcal{F}, \mathbb{P})$, $k \in \mathbb{N}$, and

$$\sum_{k, \ell=0}^{\infty} \mathbb{E}[(D_\ell u_k)^2] < \infty. \quad (3.6)$$

Then, we have

$$d_K(F, \mathcal{N}) \leq \mathbb{E}[|1 - \langle -DL^{-1}F, DF \rangle_{\ell^2(\mathbb{N})}|] + \frac{1}{2} \mathbb{E} \left[\left(|F| + \frac{\sqrt{2\pi}}{4} \right) \langle (pq)^{-1/2}, |DL^{-1}F| |DF|^2 \rangle_{\ell^2(\mathbb{N})} \right] \\ + \mathbb{E} \left[|\delta((pq)^{-1/2} DF | DL^{-1}F)|^2 \right]^{1/2}.$$

Two remarks are in order before proving Theorem 3.1 and Corollary 3.2.

Remark 3.3. *If we assume that, under \mathbb{P} , the random variables $\{\pi_n\}_{n \in \mathbb{N}}$ are independent (and so the quantities $\{p_n, q_n\}_{n \in \mathbb{N}}$ are deterministic), then the bound (3.4) coincides with that of Proposition 4.1 of [9].*

Remark 3.4. *We provide two random variables G which satisfy (3.3). By (2.16) and the fact that $\mathbb{E}[F] = 0$, one can take $G = -L^{-1}F$. Note indeed that for F satisfying the assumptions of Theorem 3.1 we have $f_x(F) \in \text{Dom}(D)$. One can also take G so that $G_n^i := \mathbb{E}[F_n^i | \mathcal{F}_{n-1}]$, $i = 0, 1$. Indeed, by Proposition 2.12,*

$$\begin{aligned} \text{Cov}(F, f_x(F)) &= \mathbb{E} \left[\sum_{n \in \mathbb{N}} \mathbb{E}[D_n F | \mathcal{F}_{n-1}] D_n f_x(F) \right] \\ &= \mathbb{E} \left[\sum_{n \in \mathbb{N}} \sqrt{p_n q_n} \mathbb{E}[F_n^1 - F_n^0 | \mathcal{F}_{n-1}] D_n f_x(F) \right] \\ &= \mathbb{E}[\langle DG, Df_x(F) \rangle_{\ell^2(\mathbb{N})}]. \end{aligned}$$

Proof of Theorem 3.1. Let f_x be the solution of the Stein equation. Due to (3.2), the claim follows if we properly bound from above the quantity $|\mathbb{E}[f'_x(F) - F f_x(F)]|$ uniformly in x . Hereon, for ease of notation, we put $f := f_x$. By the assumption (3.3), we have

$$\mathbb{E}[f'(F) - F f(F)] = \mathbb{E}[f'(F)] - \text{Cov}(F, f(F)) = \mathbb{E}[f'(F)] - \mathbb{E}[\langle DG, Df(F) \rangle_{\ell^2(\mathbb{N})}]. \quad (3.7)$$

By the definition of the gradient and the first relation in (2.6), for any $n \in \mathbb{N}$,

$$\begin{aligned} D_n f(F) &= p_n f(F_n^0) v_n^{(0)} + q_n f(F_n^1) v_n^{(1)} \\ &= p_n (f(F_n^0) - f(F)) v_n^{(0)} + q_n (f(F_n^1) - f(F)) v_n^{(1)} \\ &= p_n v_n^{(0)} \int_0^{F_n^0 - F} f'(F+t) dt + q_n v_n^{(1)} \int_0^{F_n^1 - F} f'(F+t) dt \\ &= p_n v_n^{(0)} \left(\int_0^{F_n^0 - F} (f'(F+t) - f'(F)) dt + \int_0^{F_n^0 - F} f'(F) dt \right) \\ &\quad + q_n v_n^{(1)} \left(\int_0^{F_n^1 - F} (f'(F+t) - f'(F)) dt + \int_0^{F_n^1 - F} f'(F) dt \right) \end{aligned}$$

$$= f'(F)D_n F + p_n v_n^{(0)} \int_0^{F_n^0 - F} (f'(F+t) - f'(F)) dt + q_n v_n^{(1)} \int_0^{F_n^1 - F} (f'(F+t) - f'(F)) dt.$$

By this relation and (3.7), setting for ease of notation $p_n^{(0)} := p_n$ and $p_n^{(1)} := q_n$, we have

$$\begin{aligned} \mathbb{E}[f'(F) - Ff(F)] &= \mathbb{E} \left[f'(F) \left(1 - \sum_{n \in \mathbb{N}} D_n F \left(\sum_{i=0}^1 p_n^{(i)} G_n^i v_n^{(i)} \right) \right) \right] \\ &\quad - \sum_{n \in \mathbb{N}} \mathbb{E} \left[\left(\sum_{i=0}^1 p_n^{(i)} v_n^{(i)} \int_0^{F_n^i - F} (f'(F+t) - f'(F)) dt \right) \left(\sum_{j=0}^1 p_n^{(j)} G_n^j v_n^{(j)} \right) \right]. \end{aligned}$$

By Lemma 2.1, the above rewrites as

$$\begin{aligned} \mathbb{E}[f'(F) - Ff(F)] &= \mathbb{E} \left[f'(F) \left(1 - \sum_{n \in \mathbb{N}} p_n q_n (F_n^0 - F_n^1) (G_n^0 - G_n^1) \right) \right] \\ &\quad - \sum_{n \in \mathbb{N}} \mathbb{E} \left[p_n q_n (G_n^0 - G_n^1) \int_{F_n^1 - F}^{F_n^0 - F} [f'(F+t) - f'(F)] dt \right]. \end{aligned}$$

Since f is solution of the Stein equation, it satisfies

$$f'(F+t) = (F+t)f(F+t) + \mathbf{1}_{(-\infty, x]}(F+t) - \Phi(x), \quad \forall t \in \mathbb{R}$$

and so for every $n \in \mathbb{N}$, we have

$$\begin{aligned} \int_{F_n^1 - F}^{F_n^0 - F} [f'(F+t) - f'(F)] dt &= \underbrace{\int_{F_n^1 - F}^{F_n^0 - F} [(F+t)f(F+t) - Ff(F)] dt}_{=: K_{0,1}(n)} \\ &\quad + \underbrace{\int_{F_n^1 - F}^{F_n^0 - F} [\mathbf{1}_{(-\infty, x]}(F+t) - \mathbf{1}_{(-\infty, x]}(F)] dt}_{=: L_{0,1}(n)}. \end{aligned}$$

Therefore, since $\|f'\|_\infty \leq 1$ we get

$$\begin{aligned} |\mathbb{E}[f'(F) - Ff(F)]| &\leq \mathbb{E} \left[\left| 1 - \sum_{n \in \mathbb{N}} p_n q_n (F_n^0 - F_n^1) (G_n^0 - G_n^1) \right| \right] \\ &\quad + \sum_{n \in \mathbb{N}} \mathbb{E} \left[p_n q_n |G_n^0 - G_n^1| (|K_{0,1}(n)| + |L_{0,1}(n)|) \right]. \end{aligned} \quad (3.8)$$

Now, we shall bound $|K_{0,1}(n)|$ and $|L_{0,1}(n)|$. We start by bounding $|K_{0,1}(n)|$. Applying (3.1) with $w := F$, $u := t$ and $v := 0$ yields

$$|K_{0,1}(n)| \leq \left(|F| + \frac{\sqrt{2\pi}}{4} \right) \int_{\min\{F_n^0, F_n^1\} - F}^{\max\{F_n^0, F_n^1\} - F} |t| dt. \quad (3.9)$$

In order to bound the integral appearing on the right-hand side of (3.9), we remark that for $x, y \in \mathbb{R}$ such that $y > x$ we have

$$\begin{aligned}
\int_x^y |t| dt &= \frac{1}{2} \left(\mathbf{1}_{\{x>0\}} (y^2 - x^2) + \mathbf{1}_{\{y \leq 0\}} (x^2 - y^2) + \mathbf{1}_{\{x \leq 0 < y\}} (y^2 + x^2) \right) \\
&= \frac{1}{2} \left(\mathbf{1}_{\{x>0\}} (y - x) (|x| + |y|) + \mathbf{1}_{\{y \leq 0\}} (y - x) (|x| + |y|) + \mathbf{1}_{\{x \leq 0 < y\}} (|y| \cdot y - |x| \cdot x) \right) \\
&\leq \frac{1}{2} \left(\mathbf{1}_{\{x>0\}} (y - x) (|x| + |y|) + \mathbf{1}_{\{y \leq 0\}} (y - x) (|x| + |y|) + \mathbf{1}_{\{x \leq 0 < y\}} (|y|(y - x) + |x|(y - x)) \right) \\
&= \frac{1}{2} (y - x) (|x| + |y|). \tag{3.10}
\end{aligned}$$

Taking $x := \min\{F_n^0, F_n^1\} - F$ and $y := \max\{F_n^0, F_n^1\} - F$ in (3.10) and combining the above bound with (3.9), we obtain

$$|K_{0,1}(n)| \leq \frac{1}{2} \left(|F| + \frac{\sqrt{2\pi}}{4} \right) |F_n^0 - F_n^1| (|F_n^0 - F| + |F_n^1 - F|),$$

and therefore

$$p_n q_n |G_n^0 - G_n^1| |K_{0,1}(n)| \leq \frac{1}{2} \left(|F| + \frac{\sqrt{2\pi}}{4} \right) p_n q_n |G_n^0 - G_n^1| |F_n^0 - F_n^1| (|F_n^0 - F| + |F_n^1 - F|). \tag{3.11}$$

We now provide an upper bound for $|L_{0,1}(n)|$. We have

$$\begin{aligned}
L_{0,1}(n) &= \int_{F_n^1}^{F_n^0} \mathbf{1}_{\{u \leq x\}} du - (F_n^0 - F_n^1) \mathbf{1}_{\{F \leq x\}} \\
&= (F_n^0 - F_n^1) \mathbf{1}_{\{F_n^0 \leq x, F_n^1 \leq x\}} + (F_n^0 - x) \mathbf{1}_{\{F_n^0 \leq x, F_n^1 > x\}} + (x - F_n^1) \mathbf{1}_{\{F_n^0 > x, F_n^1 \leq x\}} - (F_n^0 - F_n^1) \mathbf{1}_{\{F \leq x\}} \\
&= \mathbf{1}_{\{F \leq x\}} \left(-(F_n^0 - F_n^1) \mathbf{1}_{\{F_n^0 > x, F_n^1 > x\}} + (F_n^1 - x) \mathbf{1}_{\{F_n^0 \leq x, F_n^1 > x\}} + (x - F_n^0) \mathbf{1}_{\{F_n^0 > x, F_n^1 \leq x\}} \right) \\
&\quad + \mathbf{1}_{\{F > x\}} \left((F_n^0 - F_n^1) \mathbf{1}_{\{F_n^0 \leq x, F_n^1 \leq x\}} + (F_n^0 - x) \mathbf{1}_{\{F_n^0 \leq x, F_n^1 > x\}} + (x - F_n^1) \mathbf{1}_{\{F_n^0 > x, F_n^1 \leq x\}} \right),
\end{aligned}$$

and therefore

$$\begin{aligned}
|L_{0,1}(n)| &= \mathbf{1}_{\{F \leq x\}} \left(|F_n^0 - F_n^1| \mathbf{1}_{\{F_n^0 > x, F_n^1 > x\}} + (F_n^1 - x) \mathbf{1}_{\{F_n^0 \leq x, F_n^1 > x\}} + (F_n^0 - x) \mathbf{1}_{\{F_n^0 > x, F_n^1 \leq x\}} \right) \\
&\quad + \mathbf{1}_{\{F > x\}} \left(|F_n^0 - F_n^1| \mathbf{1}_{\{F_n^0 \leq x, F_n^1 \leq x\}} + (x - F_n^0) \mathbf{1}_{\{F_n^0 \leq x, F_n^1 > x\}} + (x - F_n^1) \mathbf{1}_{\{F_n^0 > x, F_n^1 \leq x\}} \right) \\
&\leq \mathbf{1}_{\{F \leq x\}} |F_n^0 - F_n^1| \left(\mathbf{1}_{\{F_n^0 > x, F_n^1 > x\}} + \mathbf{1}_{\{F_n^0 \leq x, F_n^1 > x\}} + \mathbf{1}_{\{F_n^0 > x, F_n^1 \leq x\}} \right) \\
&\quad + \mathbf{1}_{\{F > x\}} |F_n^0 - F_n^1| \left(\mathbf{1}_{\{F_n^0 \leq x, F_n^1 \leq x\}} + \mathbf{1}_{\{F_n^0 \leq x, F_n^1 > x\}} + \mathbf{1}_{\{F_n^0 > x, F_n^1 \leq x\}} \right) \\
&= |F_n^0 - F_n^1| \left(\mathbf{1}_{\{F > x, \min\{F_n^0, F_n^1\} \leq x\}} + \mathbf{1}_{\{F \leq x, \max\{F_n^0, F_n^1\} > x\}} \right) \\
&= |F_n^0 - F_n^1| \mathbf{1}_{\{F > x, \min\{F_n^0, F_n^1\} \leq x\} \cup \{F \leq x, \max\{F_n^0, F_n^1\} > x\}}. \tag{3.12}
\end{aligned}$$

By (3.8), (3.11) and (3.12), taking the supremum over $x \in \mathbb{R}$, we have

$$d_K(F, \mathcal{N}) \leq \mathbb{E} \left[\left| 1 - \sum_{n \in \mathbb{N}} p_n q_n (F_n^0 - F_n^1) (G_n^0 - G_n^1) \right| \right]$$

$$\begin{aligned}
& + \frac{1}{2} \sum_{n \in \mathbb{N}} \mathbb{E} \left[\left(|F| + \frac{\sqrt{2\pi}}{4} \right) p_n q_n |G_n^0 - G_n^1| |F_n^0 - F_n^1| (|F_n^0 - F| + |F_n^1 - F|) \right] \\
& + \sup_{x \in \mathbb{R}} \sum_{n \in \mathbb{N}} \mathbb{E} [p_n q_n |G_n^0 - G_n^1| |F_n^0 - F_n^1| \mathbf{1}_{\{F > x, \min\{F_n^0, F_n^1\} \leq x\}} \cup \{F \leq x, \max\{F_n^0, F_n^1\} > x\}}]. \quad (3.13)
\end{aligned}$$

We also have

$$\begin{aligned}
|F_n^0 - F| + |F_n^1 - F| &= (|F_n^0 - F| + |F_n^1 - F|) \mathbf{1}_{\{\pi_n=0\}} + (|F_n^0 - F| + |F_n^1 - F|) \mathbf{1}_{\{\pi_n=1\}} \\
&= |F_n^1 - F_n^0| \mathbf{1}_{\{\pi_n=0\}} + |F_n^0 - F_n^1| \mathbf{1}_{\{\pi_n=1\}} \\
&= |F_n^1 - F_n^0|. \quad (3.14)
\end{aligned}$$

Additionally, we note that

$$\begin{aligned}
\{F > x, \min\{F_n^0, F_n^1\} \leq x\} \\
&= \{F_n^0 > x, \min\{F_n^0, F_n^1\} \leq x, \pi_n = 0\} \cup \{F_n^1 > x, \min\{F_n^0, F_n^1\} \leq x, \pi_n = 1\} \\
&= \{F_n^0 > x, F_n^1 \leq x, \pi_n = 0\} \cup \{F_n^1 > x, F_n^0 \leq x, \pi_n = 1\} \\
&= \{\max\{F_n^0, F_n^1\} > x, \min\{F_n^0, F_n^1\} \leq x, \pi_n = 0\} \cup \{\max\{F_n^0, F_n^1\} > x, \min\{F_n^0, F_n^1\} \leq x, \pi_n = 1\} \\
&= \{\max\{F_n^0, F_n^1\} > x, \min\{F_n^0, F_n^1\} \leq x\},
\end{aligned}$$

and similarly

$$\{F \leq x, \max\{F_n^0, F_n^1\} > x\} = \{\max\{F_n^0, F_n^1\} > x, \min\{F_n^0, F_n^1\} \leq x\}.$$

Thus, by (2.10)

$$\begin{aligned}
& p_n q_n |F_n^0 - F_n^1| \mathbf{1}_{\{F > x, \min\{F_n^0, F_n^1\} \leq x\}} \cup \{F \leq x, \max\{F_n^0, F_n^1\} > x\}} \\
&= p_n q_n |F_n^0 - F_n^1| \mathbf{1}_{\{\min\{F_n^0, F_n^1\} \leq x < \max\{F_n^0, F_n^1\}\}} \\
&= p_n q_n |F_n^0 - F_n^1| (\mathbf{1}_{\{F_n^0 \leq x < F_n^1\}} + \mathbf{1}_{\{F_n^1 \leq x < F_n^0\}}) \\
&= p_n q_n (F_n^1 - F_n^0) (\mathbf{1}_{\{F_n^0 \leq x < F_n^1\}} - \mathbf{1}_{\{F_n^1 \leq x < F_n^0\}}) \\
&= p_n q_n (F_n^1 - F_n^0) (\mathbf{1}_{\{F_n^1 > x\}} (1 - \mathbf{1}_{\{F_n^0 > x\}}) - \mathbf{1}_{\{F_n^0 > x\}} (1 - \mathbf{1}_{\{F_n^1 > x\}})) \\
&= p_n q_n (F_n^1 - F_n^0) (\mathbf{1}_{\{F_n^1 > x\}} - \mathbf{1}_{\{F_n^0 > x\}}) \\
&= D_n F D_n \mathbf{1}_{\{F > x\}}. \quad (3.15)
\end{aligned}$$

The proof is concluded by plugging (2.10), (3.14) and (3.15) into (3.13).

□

Proof of Corollary 3.2. The claim follows by Theorem 3.1 if we prove

$$\mathbb{E} [\langle (pq)^{-1/2} D F D \mathbf{1}_{\{F > x\}}, |DL^{-1}F| \rangle_{\ell^2(\mathbb{N})}] \leq \sqrt{\mathbb{E} [|\delta((pq)^{-1/2} D F |DL^{-1}F|)^2]}, \quad (3.16)$$

uniformly in $x \in \mathbb{R}$.

We shall check later on that the integration by parts formula of Proposition 2.11 can be applied with $\mathbf{1}_{\{F>x\}}$ in place of F and $u_k := (p_k q_k)^{-1/2} D_k F |D_k L^{-1} F|$. We have

$$\begin{aligned} \mathbb{E}[\langle (pq)^{-1/2} D F D \mathbf{1}_{\{F>x\}}, |DL^{-1} F| \rangle_{\ell^2(\mathbb{N})}] &= \mathbb{E}[\langle D \mathbf{1}_{\{F>x\}}, (pq)^{-1/2} D F |DL^{-1} F| \rangle_{\ell^2(\mathbb{N})}] \\ &= \mathbb{E}[\mathbf{1}_{\{F>x\}} \delta((pq)^{-1/2} D F |DL^{-1} F|)] \\ &\leq \mathbb{E}[|\delta((pq)^{-1/2} D F |DL^{-1} F|)|] \\ &\leq \mathbb{E}[|\delta((pq)^{-1/2} D F |DL^{-1} F|)|^2]^{1/2}, \end{aligned}$$

which gives (3.16).

Now we check the assumptions of Proposition 2.11. We start by noticing that, for any $k \in \mathbb{N}$,

$$D_k \mathbf{1}_{\{F>x\}} u_k = (p_k q_k)^{-1/2} D_k \mathbf{1}_{\{F>x\}} D_k F |D_k L^{-1} F| \geq 0, \quad \mathbb{P}\text{-a.s.} \quad (3.17)$$

Indeed,

$$D_k \mathbf{1}_{\{F>x\}} D_k F = p_k q_k (\mathbf{1}_{\{F_k^1>x\}} - \mathbf{1}_{\{F_k^0>x\}}) (F_k^1 - F_k^0) \geq 0.$$

Note that the summability condition (2.23) is guaranteed by (3.6). Indeed, since $u_k \in L^2(\Omega, \mathcal{F}, \mathbb{P})$, we have that u_k admits a chaos decomposition. Suppose that

$$u_k = \sum_{n=0}^{\infty} J_n(g_{n+1}(*, k)).$$

Then,

$$\begin{aligned} \sum_{k, \ell=0}^{\infty} \mathbb{E}[(D_{\ell} u_k)^2] &= \sum_{k, \ell=0}^{\infty} \sum_{n=1}^{\infty} n^2 \mathbb{E}[|J_{n-1}(g_{n+1}(*, \ell, k))|^2] \\ &= \sum_{k, \ell=0}^{\infty} \sum_{n=1}^{\infty} n^2 (n-1)! \|g_{n+1}(*, \ell, k)\|_{\ell^2(\mathbb{N})^{\otimes(n-1)}}^2 \\ &= \sum_{n=1}^{\infty} n \cdot n! \|g_{n+1}\|_{\ell^2(\mathbb{N})^{\otimes(n+1)}}^2. \end{aligned}$$

This clearly implies

$$\sum_{n=2}^{\infty} (n+1)! \|g_{n+1}\|_{\ell^2(\mathbb{N})^{\otimes(n+1)}}^2 \leq 2 \sum_{k, \ell=0}^{\infty} \mathbb{E}[(D_{\ell} u_k)^2]$$

and so the summability condition (2.23) follows by the assumption (3.6).

□

4 Application to first order stochastic integrals and normal random walks

In this section we let $f_N = \{f_N(n)\}_{n \in \mathbb{N}}$, for some $N \geq 1$, and consider the first order stochastic integral

$$F_N := J_1(f_N) = \sum_{n=0}^N f_N(n) Y_n.$$

Corollary 4.1. *Assume that, for some $c > 0$,*

$$p_n q_n \geq c > 0 \quad \mathbb{P}\text{-a.s. for any } n \in \{0, \dots, N\}.$$

We have the bound

$$\begin{aligned} d_K(F_N, \mathcal{N}) &\leq \left| 1 - \sum_{n=0}^N |f_N(n)|^2 \right| \\ &\quad + \frac{1}{2\sqrt{c}} \left(\frac{\sqrt{2\pi}}{4} + \sum_{n=0}^N |f_N(n)|^2 \right) \sqrt{\sum_{n=0}^N |f_N(n)|^2} \sqrt{\sum_{n=0}^N |f_N(n)|^4} + \frac{1}{\sqrt{c}} \sqrt{\sum_{n=0}^N |f_N(n)|^4}. \end{aligned} \quad (4.1)$$

Proof. A straightforward computation shows that $D_n F_N = f_N(n)$ and $D_n L^{-1} F_N = -f_N(n)$, $n \in \mathbb{N}$. In particular,

$$u_n^{(N)} := (p_n q_n)^{-1/2} D_n F_N |D_n L^{-1} F_N| = (p_n q_n)^{-1/2} f_N(n) |f_N(n)|$$

is \mathcal{F}_{n-1} -measurable, and so by Proposition 2.10 we have

$$\mathbb{E}[\delta((pq)^{-1/2} D F_N |D L^{-1} F_N|)^2] = \sum_{n=0}^N |f_N(n)|^4 \mathbb{E}[(p_n q_n)^{-1}] \leq c^{-1} \sum_{n=0}^N |f_N(n)|^4.$$

Thus by Corollary 3.2, we have

$$\begin{aligned} d_K(F_N, \mathcal{N}) &\leq \left| 1 - \sum_{n=0}^N |f_N(n)|^2 \right| + \frac{1}{2\sqrt{c}} \sum_{n=0}^N |f_N(n)|^3 \mathbb{E} \left[|F_N| + \frac{\sqrt{2\pi}}{4} \right] + \frac{1}{\sqrt{c}} \sqrt{\sum_{n=0}^N |f_N(n)|^4} \\ &\leq \left| 1 - \sum_{n=0}^N |f_N(n)|^2 \right| + \frac{1}{2\sqrt{c}} \left(\frac{\sqrt{2\pi}}{4} + \sum_{n=0}^N |f_N(n)|^2 \right) \sum_{n=0}^N |f_N(n)|^3 + \frac{1}{\sqrt{c}} \sqrt{\sum_{n=0}^N |f_N(n)|^4} \\ &\leq \left| 1 - \sum_{n=0}^N |f_N(n)|^2 \right| + \frac{1}{2\sqrt{c}} \left(\frac{\sqrt{2\pi}}{4} + \sum_{n=0}^N |f_N(n)|^2 \right) \sqrt{\sum_{n=0}^N |f_N(n)|^2} \sqrt{\sum_{n=0}^N |f_N(n)|^4} \\ &\quad + \frac{1}{\sqrt{c}} \sqrt{\sum_{n=0}^N |f_N(n)|^4}. \end{aligned}$$

□

In particular, if

$$\sum_{n=0}^N |f_N(n)|^2 \rightarrow 1 \quad \text{and} \quad \sum_{n=0}^N |f_N(n)|^4 \rightarrow 0, \quad \text{as } N \rightarrow \infty,$$

by Corollary 4.1 we have

$$d_K(F_N, \mathcal{N}) = O \left(\max \left\{ \left| 1 - \sum_{n=0}^N |f_N(n)|^2 \right|, \sqrt{\sum_{n=0}^N |f_N(n)|^4} \right\} \right), \quad \text{as } N \rightarrow +\infty. \quad (4.2)$$

Taking $f_N(n) := 1/\sqrt{N+1}$, $n \in \{0, 1, \dots, N\}$, Corollary 4.1 yields the Berry-Esseen (non asymptotic) bound

$$d_K \left(\frac{1}{\sqrt{N+1}} \sum_{n=0}^N Y_n, \mathcal{N} \right) \leq \frac{1}{2\sqrt{c}} \left(\frac{\sqrt{2\pi}}{4} + 3 \right) \frac{1}{\sqrt{N+1}}. \quad (4.3)$$

The above bound is satisfied in particular, in the framework of the Markovian Example 2.3, provided that

$$p_0 > c, \quad q_0 > c, \quad \inf_{n,i,j} P_{i,j}^{(n)} > c,$$

for some $c \in (0, 1)$.

In the framework of Example 2.4, Corollary 4.1 shows that, as N tends to infinity, S_N converges in distribution to the lognormal random variable $S_0 e^{N^{-1/2}}$, provided that

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^N \sigma^2(n) = 1, \quad \lim_{N \rightarrow \infty} \frac{1}{N^2} \sum_{n=0}^N \sigma^4(n) = 0$$

and $p_n q_n \geq c > 0$ for any $n \in \mathbb{N}$.

References

- [1] S. Attal, J. Deschamps, and C. Pellegrini. Complex obtuse random walks and their continuous-time limits. *Probab. Theory Related Fields*, 165(1-2):65–116, 2016.
- [2] S. Attal and M. Émery. Équations de structure pour des martingales vectorielles. In *Séminaire de Probabilités, XXVIII*, volume 1583 of *Lecture Notes in Math.*, pages 256–278. Springer, Berlin, 1994.
- [3] L.H.Y. Chen, L. Goldstein, and Q.-M. Shao. *Normal approximation by Stein's method*. Probability and its Applications (New York). Springer, Heidelberg, 2011.
- [4] I. Flint, N. Privault, and G.L. Torrisi. Bounds in total variation distance for discrete-time processes on the sequence space. *Potential Anal.*, 52(2):223–243, 2020.
- [5] U. Franz and T. Hamdi. Stochastic analysis for obtuse random walks. *J. Theoret. Probab.*, 28(2):619–649, 2015.
- [6] H. Halconruy. Malliavin calculus for marked binomial processes: portfolio optimisation in the trinomial model and compound Poisson approximation. Preprint arXiv:2104.00914, 2021.
- [7] K. Krokowski. Poisson approximation of Rademacher functionals by the Chen-Stein method and Malliavin calculus. *Commun. Stoch. Anal.*, 11(2):195–222, 2017.
- [8] K. Krokowski, A. Reichenbachs, and C. Thäle. Berry-Esseen bounds and multivariate limit theorems for functionals of Rademacher sequences. *Ann. Inst. Henri Poincaré Probab. Stat.*, 52(2):763–803, 2016.
- [9] K. Krokowski, A. Reichenbachs, and C. Thäle. Discrete Malliavin-Stein method: Berry-Esseen bounds for random graphs and percolation. *Ann. Probab.*, 45(2):1071–1109, 2017.

- [10] I. Nourdin and G. Peccati. Stein's method on Wiener chaos. *Probab. Theory Related Fields*, 145(1-2):75–118, 2009.
- [11] I. Nourdin, G. Peccati, and G. Reinert. Stein's method and stochastic analysis of Rademacher functionals. *Electron. J. Probab.*, 15(55):1703–1742, 2010.
- [12] G. Peccati, J. L. Solé, M. S. Taqqu, and F. Utzet. Stein's method and normal approximation of Poisson functionals. *Ann. Probab.*, 38(2):443–478, 2010.
- [13] G. Peccati and C. Thäle. Gamma limits and U-statistics on the Poisson space. *ALEA Lat. Am. J. Probab. Math. Stat.*, 10(1):525–560, 2013.
- [14] N. Privault. *Stochastic analysis in discrete and continuous settings: with normal martingales*, volume 1982 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2009.
- [15] N. Privault and G.L. Torrisi. The Stein and Chen-Stein methods for functionals of non-symmetric Bernoulli processes. *ALEA Lat. Am. J. Probab. Math. Stat.*, 12:309–356, 2015.