

## Stein’s method and stochastic analysis of Rademacher functionals

Ivan Nourdin\*,  
Université Paris VI

Giovanni Peccati†  
Université Paris Ouest

Gesine Reinert‡  
Oxford University

### Abstract

We compute explicit bounds in the Gaussian approximation of functionals of infinite Rademacher sequences. Our tools involve Stein’s method, as well as the use of appropriate discrete Malliavin operators. As the bounds are given in terms of Malliavin operators, no coupling construction is required. When the functional depends only on the first  $d$  coordinates of the Rademacher sequence, a simple sufficient condition for convergence to a normal distribution is derived. For finite quadratic forms, we obtain necessary and sufficient conditions.

Although our approach does not require the classical use of exchangeable pairs, when the functional depends only on the first  $d$  coordinates of the Rademacher sequence we employ chaos expansion in order to construct an explicit exchangeable pair vector; the elements of the vector relate to the summands in the chaos decomposition and satisfy a linearity condition for the conditional expectation.

Among several examples, such as random variables which depend on infinitely many Rademacher variables, we provide three main applications: (i) to CLTs for multilinear forms belonging to a fixed chaos, (ii) to the Gaussian approximation of weighted infinite 2-runs, and (iii) to the computation of explicit bounds in CLTs for multiple integrals over sparse sets. This last application provides an alternate proof (and several refinements) of a recent result by Blei and Janson.

---

\*Laboratoire de Probabilités et Modèles Aléatoires, Université Pierre et Marie Curie (Paris VI), Boîte courrier 188, 4 place Jussieu, 75252 Paris Cedex 05, France. Email: [ivan.nourdin@upmc.fr](mailto:ivan.nourdin@upmc.fr)

†Equipe Modal’X, Université Paris Ouest – Nanterre la Défense, 200 Avenue de la République, 92000 Nanterre, and LSTA, Université Paris VI, France. Email: [giovanni.peccati@gmail.com](mailto:giovanni.peccati@gmail.com)

‡Department of Statistics, University of Oxford, 1 South Parks Road, Oxford OX1 3TG, UK. Email [reinert@stats.ox.ac.uk](mailto:reinert@stats.ox.ac.uk)

**Key words:** Central Limit Theorems; Discrete Malliavin operators; Normal approximation; Rademacher sequences; Sparse sets; Stein's method; Walsh chaos.

**AMS 2000 Subject Classification:** Primary 60F05; 60F99; 60G50; 60H07.

Submitted to EJP on July 6, 2009, final version accepted September 11, 2010.

# 1 Introduction

The connection between Stein's method (see e.g. [8, 36, 43, 44]) and the integration by parts formulae of stochastic analysis has been the object of a recent and quite fruitful study. The papers [25, 26, 27] deal with Stein's method and Malliavin calculus (see e.g. [29]) in a Gaussian setting; in [28] one can find extensions involving density estimates and concentration inequalities; the paper [31] is devoted to explicit bounds, obtained by combining Malliavin calculus and Stein's method in the framework of Poisson measures. Note that all these references contain examples and applications that were previously outside the scope of Stein's method, as they involve functionals of infinite-dimensional random fields for which the usual Stein-type techniques (such as exchangeable pairs, zero-bias transforms or size-bias couplings), which involve picking a random index from a *finite* set, seem inappropriate.

The aim of this paper is to push this line of research one step further, by combining Stein's method with an appropriate *discrete* version of Malliavin calculus, in order to study the normal approximation of the functionals of an infinite *Rademacher sequence*. By this expression we simply mean a sequence  $X = \{X_n : n \geq 1\}$  of i.i.d. random variables, such that  $P(X_1 = 1) = P(X_1 = -1) = 1/2$ . A full treatment of the discrete version of Malliavin calculus adopted in this paper can be found in [34] or [35]; Section 2.5 below provides a short introduction to the main objects and results.

The main features of our approach are the following:

- (i) We will be able to deal directly with the normal approximation of random variables possibly depending on the whole *infinite* sequence  $X$ . In particular, our results apply to random elements not necessarily having the form of partial sums.
- (ii) We will obtain bounds expressed in terms of the kernels appearing in the *chaotic decomposition* of a given square-integrable random variable. Note that every square-integrable functional of  $X$  admits a chaotic decomposition into an orthogonal sum of multiple integrals.
- (iii) We employ the chaotic expansion in order to construct an explicit exchangeable pair vector for any random variable which depends on a finite set of Rademacher variables. In particular, this exchangeable pair satisfies a linearity condition in conditional expectation *exactly*, which allows the normal approximation results from Reinert and Röllin [37] to be applied.
- (iv) In the particular case of random variables belonging to a *fixed chaos*, that is, having the form of a (possibly infinite) series of multilinear forms of a fixed order, we will express our bounds in terms of norms of *contraction operators* (see Section 2.2). These objects also play a central role in the normal and Gamma approximations of functionals of Gaussian fields (see [25, 26, 27, 30]) and Poisson measures (see [31, 32, 33]). In this paper, we shall consider kernels defined on discrete sets, and the corresponding contraction norms are integrals with respect to appropriate tensor products of the counting measure.

The use of contraction operators in the normal approximation of Rademacher functionals seems to us a truly promising direction for further research. One striking application of these techniques is given in Section 6, where we deduce explicit bounds (as well as a new proof) for a combinatorial central limit theorem (CLT) recently proved by Blei and Janson in [5], involving sequences of multiple integrals over finite "sparse" sets. In the original Blei and Janson's proof, the required sparseness condition emerges as an efficient way of re-expressing moment computations related to

martingale CLTs. In contrast, our approach does not involve martingales or moments and (quite surprisingly) the correct sparseness condition stems naturally from the definition of contraction. This yields moreover an explicit upper bound on an adequate distance between probability measures.

Another related work is the paper by Mossel *et al.* [24], where the authors prove a general invariance principle, allowing to compare the distribution of polynomial forms constructed from different collections of i.i.d. random variables. In particular, in [24] a crucial role is played by kernels “with low influences”: we will point out in Section 4 that influence indices are an important component of our bounds for the normal approximation of elements of a fixed chaos. It is to note that influence indices already appear (under a different name) in the papers by Rotar’ [40, 41]. In Section 6.3, we also show that the combination of the findings of the present paper with those of [24] yields an elegant answer to a question left open by Blei and Janson in [5].

A further important point is that our results do not hinge at all on the fact that  $\mathbb{N}$  is an ordered set. This is one of the advantages both of Stein’s method and of the Malliavin-type, filtration-free stochastic analysis. It follows that our findings can be directly applied to i.i.d. Rademacher families indexed by an arbitrary discrete set.

The reader is referred to Chatterjee [6] for another Stein-type approach to the normal approximation of functionals of finite i.i.d. vectors. Some connections between our findings and the results proved in [6] (in the special case of quadratic functionals) are pointed out in Section 4.2. Indeed there is a long history of normal approximations for finite vectors; notably Hoeffding decompositions, also in conjunction with truncations such as Hajek’s projection, can yield optimal Berry-Essén bounds, see [3] and [18] as well [46] for some seminal papers. A connection between jackknife and Hoeffding decompositions is explored for example in [13]. The connection between chaos decomposition and Hoeffding decompositions is detailed in Subsection 2.4. In contrast, our approach allows to directly treat functionals of infinite sequences.

The paper is organized as follows. Section 2 contains results on multiple integrals, Stein’s method and discrete Malliavin calculus (in particular, a new chain rule is proved). Section 3 provides general bounds for the normal approximation of Rademacher functionals; it gives first examples and the exchangeable pair construction. In Section 4 we are interested in random variables belonging to a fixed chaos. Section 5 focuses on sums of single and double integrals and related examples. Section 6 is devoted to further applications and refinements, mainly involving integrals over sparse sets. Technical proofs are deferred to the Appendix.

## 2 Framework and main tools

### 2.1 The setup

As before, we shall denote by  $X = \{X_n : n \geq 1\}$  a standard *Rademacher sequence*. This means that  $X$  is a collection of i.i.d. random variables, defined on a probability space  $(\Omega, \mathcal{F}, P)$  and such that  $P[X_1 = 1] = P[X_1 = -1] = 1/2$ . To simplify the subsequent discussion, for the rest of the paper we shall suppose that  $\Omega = \{-1, 1\}^{\mathbb{N}}$  and  $P = [\frac{1}{2}\{\delta_{-1} + \delta_1\}]^{\mathbb{N}}$ . Also,  $X$  will be defined as the canonical process, that is: for every  $\omega = (\omega_1, \omega_2, \dots) \in \Omega$ ,  $X_n(\omega) = \omega_n$ . Starting from  $X$ , for every  $N \geq 1$  one can build a random signed measure  $\mu_{(X,N)}$  on  $\{1, \dots, N\}$ , defined for every  $A \subset \{1, \dots, N\}$  as follows:

$$\mu_{(X,N)}(A) = \sum_{j \in A} X_j.$$

It is clear that  $\sigma\{X\} = \sigma\{\mu_{(X,N)} : N \geq 1\}$ . We shall sometimes omit the index  $N$  and write  $\mu_{(X,N)} = \mu_X$ , whenever the domain of integration is clear from the context.

We define the set  $D = \{(i_1, i_2) \in \mathbb{N}^2 : i_1 = i_2\}$  to be the *diagonal* of  $\mathbb{N}^2$ . For  $n \geq 2$ , we put

$$\Delta_n = \{(i_1, \dots, i_n) \in \mathbb{N}^n : \text{the } i_j\text{'s are all different}\}, \quad (2.1)$$

and, for  $N, n \geq 2$ ,

$$\Delta_n^N = \Delta_n \cap \{1, \dots, N\}^n \quad (2.2)$$

(so that  $\Delta_n^N = \emptyset$  if  $N < n$ ). The random signed measures  $\mu_{(X,N)}$  satisfy that, for every  $A, B \subset \{1, \dots, N\}$ ,

$$\mu_{(X,N)}^{\otimes 2}([A \times B] \cap D) = \sum_j X_j^2 \mathbf{1}_{\{j \in A\}} \mathbf{1}_{\{j \in B\}} = \kappa(A \cap B) = \#\{j : j \in A \cap B\}, \quad (2.3)$$

where  $\kappa$  is the counting measure on  $\mathbb{N}$ . The application  $A \mapsto \mu_{(X,N)}^{\otimes 2}([A \times A] \cap D)$  is called the *diagonal measure* associated with  $\mu_{(X,N)}$ . The above discussion simply says that, for every  $N \geq 1$ , the diagonal measure associated with  $\mu_{(X,N)}$  is the restriction on  $\{1, \dots, N\}$  of the counting measure  $\kappa$ . Note that, for finite sets  $A, B$  one has also

$$E[\mu_X(A)\mu_X(B)] = \kappa(A \cap B).$$

**Remark 2.1.** In the terminology e.g. of [33] and [45], the collection of random variables

$$\{\mu_{(X,N)}(A) : A \subset \{1, \dots, N\}, N \geq 1\}$$

defines an *independently scattered random measure* (also called a *completely random measure*) on  $\mathbb{N}$ , with *control measure* equal to  $\kappa$ . In particular, if  $A_1, \dots, A_k$  is a collection of mutually disjoint finite subsets of  $\mathbb{N}$ , then the random variables  $\mu_X(A_1), \dots, \mu_X(A_k)$  are mutually independent.

## 2.2 The star notation

For  $n \geq 1$ , we denote by  $\ell^2(\mathbb{N})^n$  the class of the kernels (= functions) on  $\mathbb{N}^n$  that are square integrable with respect to  $\kappa^{\otimes n}$ ;  $\ell^2(\mathbb{N})^{on}$  is the subset of  $\ell^2(\mathbb{N})^n$  composed of symmetric kernels;  $\ell_0^2(\mathbb{N})^n$  is the subset of  $\ell^2(\mathbb{N})^n$  composed of kernels vanishing on diagonals, that is, vanishing on the complement of  $\Delta_n$ ;  $\ell_0^2(\mathbb{N})^{on}$  is the subset of  $\ell_0^2(\mathbb{N})^n$  composed of symmetric kernels.

For every  $n, m \geq 1$ , every  $r = 0, \dots, n \wedge m$ , every  $l = 0, \dots, r$ , and every  $f \in \ell_0^2(\mathbb{N})^{on}$  and  $g \in \ell_0^2(\mathbb{N})^{om}$ , we denote by  $f \star_r^l g$  the function of  $n + m - r - l$  variables obtained as follows:  $r$  variables are identified and, among these,  $l$  are integrated out with respect to the counting measure  $\kappa$ . Explicitly,

$$\begin{aligned} & f \star_r^l g(i_1, \dots, i_{n+m-r-l}) \\ &= \sum_{a_1, \dots, a_l} f(i_1, \dots, i_{n-r}; i_{n-r+1}, \dots, i_{n-l}; a_1, \dots, a_l) \\ & \quad \times g(i_{n-l+1}, \dots, i_{n+m-r-l}; i_{n-r+1}, \dots, i_{n-l}; a_1, \dots, a_l); \end{aligned}$$

note that, since  $f$  and  $g$  vanish on diagonals by assumption, the sum actually runs over all vectors  $(a_1, \dots, a_l) \in \Delta_l$ . For instance,

$$f \star_0^0 g(i_1, \dots, i_{n+m}) = f \otimes g(i_1, \dots, i_{n+m}) = f(i_1, \dots, i_n) g(i_{n+1}, \dots, i_{n+m}),$$

and, for  $r \in \{1, \dots, n \wedge m\}$ ,

$$\begin{aligned} & f \star_r^f g (i_1, \dots, i_{n+m-2r}) \\ &= \sum_{a_1, \dots, a_r} f (i_1, \dots, i_{n-r}; a_1, \dots, a_r) g (i_{n-r+1}, \dots, i_{n+m-2r}; a_1, \dots, a_r). \end{aligned}$$

In particular, if  $r = m = n$

$$f \star_m^m g = \sum_{a_1, \dots, a_m} f (a_1, \dots, a_m) g (a_1, \dots, a_m) = \langle f, g \rangle_{\ell^2(\mathbb{N})^{\otimes m}}.$$

**Example 2.2.** If  $n = m = 2$ , one has

$$\begin{aligned} f \star_0^0 g (i_1, i_2, i_3, i_4) &= f (i_1, i_2) g (i_3, i_4) ; f \star_1^0 g (i_1, i_2, i_3) = f (i_1, i_2) g (i_3, i_2) ; \\ f \star_1^1 g (i_1, i_2) &= \Sigma_a f (i_1, a) g (i_2, a) ; f \star_2^0 g (i_1, i_2) = f (i_1, i_2) g (i_1, i_2) ; \\ f \star_2^1 g (i) &= \Sigma_a f (i, a) g (i, a) ; f \star_2^2 g = \Sigma_{a_1, a_2} f (a_1, a_2) g (a_1, a_2). \end{aligned}$$

**Remark 2.3.** In general, a kernel of the type  $f \star_r^l g$  may be not symmetric and may not vanish on diagonal sets.

The next two lemmas collect some useful facts for subsequent sections. Their proofs are postponed in the appendix.

**Lemma 2.4.** Let  $f \in \ell_0^2(\mathbb{N})^{\otimes n}$ ,  $g \in \ell_0^2(\mathbb{N})^{\otimes m}$  ( $n, m \geq 1$ ) and  $0 \leq l \leq r \leq n \wedge m$ . Then:

1.  $f \star_r^l g \in \ell^2(\mathbb{N})^{\otimes(n+m-r-l)}$  and  $\|f \star_r^l g\|_{\ell^2(\mathbb{N})^{\otimes(n+m-r-l)}} \leq \|f\|_{\ell^2(\mathbb{N})^{\otimes n}} \|g\|_{\ell^2(\mathbb{N})^{\otimes m}}$ ;
2. if  $n \geq 2$  then

$$\begin{aligned} \max_{j \in \mathbb{N}} \left[ \sum_{(b_1, \dots, b_{n-1}) \in \mathbb{N}^{n-1}} f^2(j, b_1, \dots, b_{n-1}) \right]^2 &\leq \|f \star_n^{n-1} f\|_{\ell^2(\mathbb{N})}^2 \\ &\leq \|f\|_{\ell^2(\mathbb{N})^{\otimes n}}^2 \times \max_{j \in \mathbb{N}} \sum_{(b_1, \dots, b_{n-1}) \in \mathbb{N}^{n-1}} f^2(j, b_1, \dots, b_{n-1}) \end{aligned}$$

and, if  $l = 1, \dots, n$ ,

$$\begin{aligned} \|f \star_l^{l-1} g\|_{\ell^2(\mathbb{N})^{\otimes(n+m-2l+1)}}^2 &= \sum_{j=1}^{\infty} \|f(j, \cdot) \star_{l-1}^{l-1} g(j, \cdot)\|_{\ell^2(\mathbb{N})^{\otimes(n+m-2l)}}^2 \\ &\leq \|f \star_n^{n-1} f\|_{\ell^2(\mathbb{N})} \times \|g\|_{\ell^2(\mathbb{N})^{\otimes m}}^2. \end{aligned} \tag{2.4}$$

3.  $\|f \star_1^0 f\|_{\ell^2(\mathbb{N})^{\otimes(2n-1)}} = \|f \star_n^{n-1} f\|_{\ell^2(\mathbb{N})}$  and, for every  $l = 2, \dots, n$  ( $n \geq 2$ ),

$$\|f \star_l^{l-1} f\|_{\ell^2(\mathbb{N})^{\otimes(2n-2l+1)}} \leq \|f \star_{l-1}^{l-1} f \times \mathbf{1}_{\Delta_{2(n-l+1)}^c}\|_{\ell^2(\mathbb{N})^{\otimes(2n-2l+2)}} \leq \|f \star_{l-1}^{l-1} f\|_{\ell^2(\mathbb{N})^{\otimes(2n-2l+2)}}.$$

Here,  $\Delta_q^c$  stands for the complement of the set  $\Delta_q$ , that is,  $\Delta_q^c$  is the collection of all vectors  $(i_1, \dots, i_q)$  such that  $i_k = i_l$  for at least one pair  $(l, k)$  such that  $l \neq k$ .

**Remark 2.5.** 1. According e.g. to [24], the quantity

$$\sum_{(b_1, \dots, b_{n-1}) \in \mathbb{N}^{n-1}} f^2(j, b_1, \dots, b_{n-1})$$

is called the *influence of the  $j^{\text{th}}$  coordinate* on  $f$ .

2. When specializing Lemma 2.4 to the case  $n = 2$ , one gets

$$\begin{aligned} \max_{j \in \mathbb{N}} \left[ \sum_{i \in \mathbb{N}} f^2(i, j) \right]^2 &\leq \|f \star_2^1 f\|_{\ell^2(\mathbb{N})}^2 \\ &= \|f \star_1^1 f \times \mathbf{1}_{\Delta_2^c}\|_{\ell^2(\mathbb{N})^{\otimes 2}}^2 \leq \|f \star_1^1 f\|_{\ell^2(\mathbb{N})^{\otimes 2}}^2 = \mathbf{Trace}\{[f]^4\}. \end{aligned} \quad (2.5)$$

Here,  $[f]$  denote the infinite array  $\{f(i, j) : i, j \in \mathbb{N}\}$  and  $[f]^4$  stands for the fourth power of  $[f]$  (regarded as the kernel of a Hilbert-Schmidt operator).

**Lemma 2.6.** Fix  $n, m \geq 1$ , and let  $f_k \in \ell_0^2(\mathbb{N})^{on}$  and  $g_k \in \ell_0^2(\mathbb{N})^{om}$ ,  $k \geq 1$ , be such that  $f_k \xrightarrow[k \rightarrow \infty]{} f$  and  $g_k \xrightarrow[k \rightarrow \infty]{} g$ , respectively, in  $\ell_0^2(\mathbb{N})^{on}$  and  $\ell_0^2(\mathbb{N})^{om}$ . Then, for every  $r = 0, \dots, n \wedge m$ ,  $f_k \star_r^r g_k \xrightarrow[k \rightarrow \infty]{} f \star_r^r g$  in  $\ell^2(\mathbb{N})^{\otimes m+n-2r}$ .

### 2.3 Multiple integrals, chaos and product formulae

Following e.g. [35], for every  $q \geq 1$  and every  $f \in \ell_0^2(\mathbb{N})^{oq}$  we denote by  $J_q(f)$  the *multiple integral* (of order  $q$ ) of  $f$  with respect to  $X$ . Explicitly,

$$\begin{aligned} J_q(f) &= \sum_{(i_1, \dots, i_q) \in \mathbb{N}^q} f(i_1, \dots, i_q) X_{i_1} \cdots X_{i_q} = \sum_{(i_1, \dots, i_q) \in \Delta^q} f(i_1, \dots, i_q) X_{i_1} \cdots X_{i_q} \\ &= q! \sum_{i_1 < \dots < i_q} f(i_1, \dots, i_q) X_{i_1} \cdots X_{i_q}, \end{aligned} \quad (2.6)$$

where the possibly infinite sum converges in  $L^2(\Omega)$ . Note that, if  $f$  has support contained in some finite set  $\{1, \dots, N\}^q$ , then  $J_q(f)$  is a true integral with respect to the product signed measure  $\mu_{(X, N)}^{\otimes q}$ , that is,

$$J_q(f) = \int_{\{1, \dots, N\}^q} f d\mu_{(X, N)}^{\otimes q} = \int_{\Delta_q^N} f d\mu_{(X, N)}^{\otimes q}.$$

One customarily sets  $\ell^2(\mathbb{N})^{o0} = \mathbb{R}$ , and  $J_0(c) = c$ ,  $\forall c \in \mathbb{R}$ . It is known (see again [35]) that, if  $f \in \ell_0^2(\mathbb{N})^{oq}$  and  $g \in \ell_0^2(\mathbb{N})^{op}$ , then one has the isometric relation

$$E[J_q(f)J_p(g)] = \mathbf{1}_{\{q=p\}} q! \langle f, g \rangle_{\ell^2(\mathbb{N})^{\otimes q}}. \quad (2.7)$$

In particular,

$$E[J_q(f)^2] = q! \|f\|_{\ell^2(\mathbb{N})^{\otimes q}}^2. \quad (2.8)$$

The collection of all random variables of the type  $J_n(f)$ , where  $f \in \ell_0^2(\mathbb{N})^{oq}$ , is called the  *$q$ th chaos* associated with  $X$ .

**Remark 2.7.** Popular names for random variables of the type  $J_q(f)$  are *Walsh chaos* (see e.g. [22, Ch. IV]) and *Rademacher chaos* (see e.g. [5, 12, 21]). In Meyer's monograph [23] the collection  $\{J_q(f) : f \in \ell_0^2(\mathbb{N})^{oq}, q \geq 0\}$  is called the *toy Fock space* associated with  $X$ .

Recall (see [35]) that one has the following *chaotic decomposition*: for every  $F \in L^2(\sigma\{X\})$  (that is, for every square integrable functional of the sequence  $X$ ) there exists a unique sequence of kernels  $f_n \in \ell_0^2(\mathbb{N})^{on}, n \geq 1$ , such that

$$F = E(F) + \sum_{n \geq 1} J_n(f_n) = E(F) + \sum_{n \geq 1} n! \sum_{i_1 < i_2 < \dots < i_n} f_n(i_1, \dots, i_n) X_{i_1} \cdots X_{i_n}, \quad (2.9)$$

where the second equality follows from the definition of  $J_n(f_n)$ , and the series converge in  $L^2$ .

**Remark 2.8.** Relation (2.9) is equivalent to the statement that the set

$$\{1\} \cup \bigcup_{n \geq 1} \{X_{i_1} \dots X_{i_n} : 1 \leq i_1 < \dots < i_n\} \quad (2.10)$$

is an orthonormal basis of  $L^2(\sigma\{X\})$ . An immediate proof of this fact can be deduced from basic harmonic analysis. Indeed, it is well-known that the set  $\Omega = \{-1, +1\}^{\mathbb{N}}$ , when endowed with the product structure of coordinate multiplication, is a compact Abelian group (known as the *Cantor group*), whose unique normalized Haar measure is the law of  $X$ . Relation (2.9) then follows from the fact that the dual of  $\Omega$  consists exactly in the mappings  $\omega \mapsto 1$  and  $\omega \mapsto X_{i_1}(\omega) \cdots X_{i_n}(\omega)$ ,  $i_n > \dots > i_1 \geq 1, n \geq 1$ . See e.g. [4, Section VII.2].

Given a kernel  $f$  on  $\mathbb{N}^n$ , we denote by  $\tilde{f}$  its canonical symmetrization, that is:

$$\tilde{f}(i_1, \dots, i_n) = \frac{1}{n!} \sum_{\sigma} f(i_{\sigma(1)}, \dots, i_{\sigma(n)}),$$

where  $\sigma$  runs over the  $n!$  permutations of  $\{1, \dots, n\}$ . The following result is a standard multiplication formula between multiple integrals of possibly different orders. Since we did not find this result in the literature (see, however, [35, formula (15)]), we provide a simple combinatorial proof in Section 6.3.

**Proposition 2.9.** For every  $n, m \geq 1$ , every  $f \in \ell_0^2(\mathbb{N})^{on}$  and  $g \in \ell_0^2(\mathbb{N})^{om}$ , one has that

$$J_n(f) J_m(g) = \sum_{r=0}^{n \wedge m} r! \binom{n}{r} \binom{m}{r} J_{n+m-2r} \left[ \left( \widetilde{f \star_r^r g} \right) \mathbf{1}_{\Delta_{n+m-2r}} \right]. \quad (2.11)$$

**Remark 2.10.** Proposition 2.9 yields that the product of two multiple integrals is a linear combination of square-integrable random variables. By induction, this implies in particular that, for every  $f \in \ell_0^2(\mathbb{N})^{on}$  and every  $k \geq 1$ ,  $E|J_n(f)|^k < \infty$ , that is, the elements belonging to a given chaos have finite moments of all orders. This fact can be alternatively deduced from standard hypercontractive inequalities, see e.g. Theorem 3.2.1 in [12] or [21, Ch. 6]. Note that similar results hold for random variables belonging to a fixed Wiener chaos of a Gaussian field (see e.g. [19, Ch. V]).



## 2.4 Finding a chaotic decomposition

The second equality in (2.9) clearly implies that, for every  $i_1 < \dots < i_n$ ,

$$n!f_n(i_1, \dots, i_n) = E(F \times X_{i_1} \cdots X_{i_n}).$$

In what follows, we will point out two alternative procedures yielding an explicit expression for the kernels  $f_n$ .

(i) *Möbius inversion (Hoeffding decompositions)*. We start with the following observation: if  $F = F(X_1, \dots, X_d)$  is a random variable that depends uniquely on the first  $d$  coordinates of the Rademacher sequence  $X$ , then necessarily  $F$  admits a finite chaotic decomposition of the form

$$F = E(F) + \sum_{n=1}^d \sum_{1 \leq i_1 < \dots < i_n \leq d} n!f_n(i_1, \dots, i_n)X_{i_1} \cdots X_{i_n}. \quad (2.12)$$

Note that the sum in the chaotic decomposition (2.12) must stop at  $d$ : indeed, the terms of order greater than  $d$  are equal to zero since, if not, they would involve e.g. products of the type  $X_{j_1} \cdots X_{j_{d+1}}$  with all the  $j_a$ 's different, which would not be consistent with the fact that  $F$  is  $\sigma\{X_1, \dots, X_d\}$ -measurable. Now note that one can rewrite (2.12) as  $F = E(F) + \sum_{I \subset \{1, \dots, d\}} G(I)$ , where, for  $I := \{i_1, \dots, i_a\}$ ,

$$G(I) = a!f_a(i_1, \dots, i_a)X_{i_1} \cdots X_{i_a}.$$

By exploiting the independence, one obtains also that, for every  $J := \{j_1, \dots, j_n\} \subset \{1, \dots, d\}$ ,

$$H(J) := E[F - E(F)|X_{j_1}, \dots, X_{j_n}] = \sum_{I \subset J} G(I),$$

thus yielding that, by inversion (see e.g. [42, p. 116]), for every  $n = 1, \dots, d$ ,

$$n!f_n(j_1, \dots, j_n)X_{j_1} \cdots X_{j_n} = \sum_{\{i_1, \dots, i_a\} \subset \{j_1, \dots, j_n\}} (-1)^{n-a} E[F - E(F)|X_{i_1}, \dots, X_{i_a}]. \quad (2.13)$$

Formula (2.13) simply means that, for a fixed  $n$ , the sum of all random variables of the type  $n!f_n(j_1, \dots, j_n)X_{j_1} \cdots X_{j_n}$  coincides with the  $n^{\text{th}}$  term in the *Hoeffding-ANOVA* decomposition of  $F$  (see e.g. [20]). By a density argument (which is left to the reader), representation (2.13) extends indeed to random variables depending on the *whole* sequence  $X$ .

(ii) *Indicators expansions*. Assume that we can write  $F = F(X_1, \dots, X_d)$  with  $F : \{-1, +1\}^d \rightarrow \mathbb{R}$ . Taking into account all the possibilities, we have

$$F = \sum_{(\varepsilon_1, \dots, \varepsilon_d) \in \{-1, +1\}^d} F(\varepsilon_1, \dots, \varepsilon_d) \prod_{i=1}^d \mathbf{1}_{\{X_i = \varepsilon_i\}}.$$

Now, the crucial point is the identity  $\mathbf{1}_{\{X_i = \varepsilon_i\}} = \frac{1}{2}(1 + \varepsilon_i X_i)$ . Hence

$$F = 2^{-d} \sum_{(\varepsilon_1, \dots, \varepsilon_d) \in \{-1, +1\}^d} F(\varepsilon_1, \dots, \varepsilon_d) \prod_{i=1}^d (1 + \varepsilon_i X_i). \quad (2.14)$$

But

$$\begin{aligned} \prod_{i=1}^d (1 + \varepsilon_i X_i) &= 1 + \sum_{1 \leq i_1 \leq d} \varepsilon_{i_1} X_{i_1} + \sum_{1 \leq i_1 < i_2 \leq d} \varepsilon_{i_1} \varepsilon_{i_2} X_{i_1} X_{i_2} \\ &+ \sum_{1 \leq i_1 < i_2 < i_3 \leq d} \varepsilon_{i_1} \varepsilon_{i_2} \varepsilon_{i_3} X_{i_1} X_{i_2} X_{i_3} + \dots + \varepsilon_{i_1} \dots \varepsilon_{i_d} X_{i_1} \dots X_{i_d}; \end{aligned}$$

inserting this in (2.14) one can deduce the chaotic expansion of  $F$ .

## 2.5 Discrete Malliavin calculus and a new chain rule

We will now define a set of discrete operators which are the analogues of the classic Gaussian-based Malliavin operators (see e.g. [19, 29]). The reader is referred to [34] and [35] for any unexplained notion and/or result.

The operator  $D$ , called the *gradient operator*, transforms random variables into random sequences. Its domain, noted  $\text{dom}D$ , is given by the class of random variables  $F \in L^2(\sigma\{X\})$  such that the kernels  $f_n \in \ell_0^2(\mathbb{N})^{\otimes n}$  in the chaotic expansion  $F = E(F) + \sum_{n \geq 1} J_n(f_n)$  (see (2.9)) verify the relation

$$\sum_{n \geq 1} n n! \|f_n\|_{\ell_0^2(\mathbb{N})^{\otimes n}}^2 < \infty.$$

In particular, if  $F = F(X_1, \dots, X_d)$  depends uniquely on the first  $d$  coordinates of  $X$ , then  $F \in \text{dom}D$ . More precisely,  $D$  is an operator with values in  $L^2(\Omega \times \mathbb{N}, P \otimes \kappa)$ , such that, for every  $F = E(F) + \sum_{n \geq 1} J_n(f_n) \in \text{dom}D$ ,

$$D_k F = \sum_{n \geq 1} n J_{n-1}(f_n(\cdot, k)), \quad k \geq 1, \quad (2.15)$$

where the symbol  $f_n(\cdot, k)$  indicates that the integration is performed with respect to  $n - 1$  variables. According e.g. to [34, 35], the gradient operator admits the following representation. Let  $\omega = (\omega_1, \omega_2, \dots) \in \Omega$ , and set

$$\omega_+^k = (\omega_1, \omega_2, \dots, \omega_{k-1}, +1, \omega_{k+1}, \dots)$$

and

$$\omega_-^k = (\omega_1, \omega_2, \dots, \omega_{k-1}, -1, \omega_{k+1}, \dots)$$

to be the sequences obtained by replacing the  $k^{\text{th}}$  coordinate of  $\omega$ , respectively, with  $+1$  and  $-1$ . Write  $F_k^\pm$  instead of  $F(\omega_\pm^k)$  for simplicity. Then, for every  $F \in \text{dom}D$ ,

$$D_k F(\omega) = \frac{1}{2} (F_k^+ - F_k^-), \quad k \geq 1. \quad (2.16)$$

**Remark 2.11.** 1. It is easily seen that, if the random variable  $F \in L^2(\sigma\{X\})$  is such that the mapping  $(\omega, k) \mapsto \frac{1}{2}(F_k^+ - F_k^-)(\omega)$  is an element of  $L^2(\Omega \times \mathbb{N}, P \otimes \kappa)$ , then necessarily  $F \in \text{dom}D$ .

2. If  $F = F(X_1, \dots, X_d)$  depends uniquely on the first  $d$  coordinates of  $X$ , then the *randomized derivative*  $\Delta_j F(X)$  of  $F$ , defined in [6], relates to  $D_k F(\omega)$  as follows. Let  $X' = (X'_1, \dots, X'_d)$  be an independent copy of  $(X_1, \dots, X_d)$ , then, for  $j = 1, \dots, d$ ,

$$\begin{aligned} \Delta_j F(X_1, \dots, X_d) &= (F_j^+ - F_j^-) \left( \mathbf{1}_{\{X_j=1, X'_j=-1\}} - \mathbf{1}_{\{X_j=-1, X'_j=1\}} \right) \\ &= 2D_j F(X_1, \dots, X_d) \left( \mathbf{1}_{\{X_j=1, X'_j=-1\}} - \mathbf{1}_{\{X_j=-1, X'_j=1\}} \right). \end{aligned} \quad (2.17)$$

A key advantage of the operator  $D_j$ , compared to  $\Delta_j$ , is that no coupling construction is required, which makes  $D_j$  easily amenable to Malliavin calculus and provides a unified treatment of normal approximations both for functionals acting on sequences of continuous random variables and on sequences of discrete random variables.

We write  $\delta$  for the adjoint of  $D$ , also called the *divergence operator*. The domain of  $\delta$  is denoted by  $\text{dom}\delta$ , and is such that  $\text{dom}\delta \subset L^2(\Omega \times \mathbb{N}, P \otimes \kappa)$ . Recall that  $\delta$  is defined via the following *integration by parts formula*: for every  $F \in \text{dom}D$  and every  $u \in \text{dom}\delta$

$$E[F\delta(u)] = E[\langle DF, u \rangle_{\ell^2(\mathbb{N})}] = \langle DF, u \rangle_{L^2(\Omega \times \mathbb{N}, P \otimes \kappa)}. \quad (2.18)$$

Now set  $L_0^2(\sigma\{X\})$  to be the subspace of  $L^2(\sigma\{X\})$  composed of centered random variables. We write  $L : L^2(\sigma\{X\}) \rightarrow L_0^2(\sigma\{X\})$  for the *Ornstein-Uhlenbeck operator*, which is defined as follows. The domain  $\text{dom}L$  of  $L$  is composed of random variables  $F = E(F) + \sum_{n \geq 1} J_n(f_n) \in L^2(\sigma\{X\})$  such that

$$\sum_{n \geq 1} n^2 n! \|f_n\|_{\ell^2(\mathbb{N})^{\otimes n}}^2 < \infty,$$

and, for  $F \in \text{dom}L$ ,

$$LF = - \sum_{n \geq 1} n J_n(f_n). \quad (2.19)$$

With respect to [34, 35], note that we choose to add a minus in the right-hand side of (2.19), in order to facilitate the connection with the paper [25]. One crucial relation between the operators  $\delta$ ,  $D$  and  $L$  is that

$$\delta D = -L. \quad (2.20)$$

The inverse of  $L$ , noted  $L^{-1}$ , is defined on  $F \in L_0^2(\sigma\{X\})$ , and is given by

$$L^{-1}F = - \sum_{n \geq 1} \frac{1}{n} J_n(f_n). \quad (2.21)$$

**Lemma 2.12.** *Let  $F \in \text{dom}D$  be centered, and  $f : \mathbb{R} \rightarrow \mathbb{R}$  be such that  $f(F) \in \text{dom}D$ . Then  $E[Ff(F)] = E[\langle Df(F), -DL^{-1}F \rangle_{\ell^2(\mathbb{N})}]$ .*

*Proof.* Using (2.20) and (2.18) consecutively, we can write  $E[Ff(F)] = E[LL^{-1}Ff(F)] = -E[\delta DL^{-1}Ff(F)] = E[\langle Df(F), -DL^{-1}F \rangle_{\ell^2(\mathbb{N})}]$ .  $\square$

Finally, we define  $\{P_t : t \geq 0\} = \{e^{tL} : t \geq 0\}$  to be the the *semi-group* associated with  $L$ , that is,

$$P_t F = \sum_{n=0}^{\infty} e^{-nt} J_n(f_n), \quad t \geq 0, \quad \text{for } F = E(F) + \sum_{n=1}^{\infty} J_n(f_n) \in L^2(\sigma\{X\}). \quad (2.22)$$

The next result will be useful throughout the paper.

**Lemma 2.13.** *Let  $F \in \text{dom}D$  and fix  $k \in \mathbb{N}$ . Then:*

1. *The random variables  $D_k F$ ,  $D_k L^{-1} F$ ,  $F_k^+$  and  $F_k^-$  are independent of  $X_k$ .*
2. *It holds that  $|F_k^+ - F| \leq 2|D_k F|$  and  $|F_k^- - F| \leq 2|D_k F|$ ,  $P$ -almost surely.*
3. *If  $F$  has zero mean, then  $E\|DL^{-1}F\|_{\ell^2(\mathbb{N})}^2 \leq E\|DF\|_{\ell^2(\mathbb{N})}^2$  with equality if and only if  $F$  is an element of the first chaos.*

*Proof.* 1. One only needs to combine the definition of  $F_k^\pm$  with (2.16).

2. Use  $F_k^\pm - F = \pm(F_k^+ - F_k^-)\mathbf{1}_{\{X_k = \mp 1\}} = \pm 2D_k F \mathbf{1}_{\{X_k = \mp 1\}}$ .

3. Let us consider the chaotic expansion of  $F$ :

$$F = \sum_{n \geq 1} J_n(f_n).$$

Then  $-D_k L^{-1} F = \sum_{n \geq 1} J_{n-1}(f_n(\cdot, k))$  and  $D_k F = \sum_{n \geq 1} n J_{n-1}(f_n(\cdot, k))$ . Therefore, using the isometric relation (2.7),

$$\begin{aligned} E\|DL^{-1}F\|_{\ell^2(\mathbb{N})}^2 &= E \sum_{k \in \mathbb{N}} \left( \sum_{n \geq 1} J_{n-1}(f_n(\cdot, k)) \right)^2 \\ &= \sum_{n \geq 1} (n-1)! \|f_n\|_{\ell^2(\mathbb{N})^{\otimes n}}^2 \\ &\leq \sum_{n \geq 1} n^2 (n-1)! \|f_n\|_{\ell^2(\mathbb{N})^{\otimes n}}^2 \\ &= E \sum_{k \in \mathbb{N}} \left( \sum_{n \geq 1} n J_{n-1}(f_n(\cdot, k)) \right)^2 = E\|DF\|_{\ell^2(\mathbb{N})}^2. \end{aligned}$$

Moreover, the previous equality shows that we have equality if and only if  $f_n = 0$  for all  $n \geq 2$ , that is, if and only if  $F$  is an element of the first chaos.  $\square$

We conclude this section by proving a chain rule involving deterministic functions of random variables in the domain of  $D$ . It should be compared with the classic chain rule of the Gaussian-based Malliavin calculus (see e.g. [29, Prop. 1.2.2]).

**Proposition 2.14.** (Chain Rule). *Let  $F \in \text{dom}D$  and  $f : \mathbb{R} \rightarrow \mathbb{R}$  be thrice differentiable with bounded third derivative. Assume moreover that  $f(F) \in \text{dom}D$ . Then, for any integer  $k$ ,  $P$ -a.s.:*

$$\left| D_k f(F) - f'(F) D_k F + \frac{1}{2} (f''(F_k^+) + f''(F_k^-)) (D_k F)^2 X_k \right| \leq \frac{10}{3} |f'''|_\infty |D_k F|^3.$$

*Proof.* By a standard Taylor expansion,

$$\begin{aligned}
D_k f(F) &= \frac{1}{2}(f(F_k^+) - f(F_k^-)) = \frac{1}{2}(f(F_k^+) - f(F)) - \frac{1}{2}(f(F_k^-) - f(F)) \\
&= \frac{1}{2}f'(F)(F_k^+ - F) + \frac{1}{4}f''(F)(F_k^+ - F)^2 + R_1 \\
&\quad - \frac{1}{2}f'(F)(F_k^- - F) - \frac{1}{4}f''(F)(F_k^- - F)^2 + R_2 \\
&= f'(F)D_k F + \\
&\quad \frac{1}{8}(f''(F_k^+) + f''(F_k^-))((F_k^+ - F)^2 - (F_k^- - F)^2) + R_1 + R_2 + R_3 \\
&= f'(F)D_k F - \frac{1}{2}(f''(F_k^+) + f''(F_k^-))(D_k F)^2 X_k + R_1 + R_2 + R_3,
\end{aligned}$$

where, using Lemma 2.13,

$$\begin{aligned}
|R_1| &\leq \frac{1}{12}|f'''|_\infty |F_k^+ - F|^3 \leq \frac{2}{3}|f'''|_\infty |D_k F|^3 \\
|R_2| &\leq \frac{1}{12}|f'''|_\infty |F_k^- - F|^3 \leq \frac{2}{3}|f'''|_\infty |D_k F|^3 \\
|R_3| &\leq \frac{1}{8}|f'''|_\infty (|F_k^+ - F|^3 + |F_k^- - F|^3) \leq 2|f'''|_\infty |D_k F|^3.
\end{aligned}$$

By putting these three inequalities together, the desired conclusion follows.  $\square$

## 2.6 Stein's method for normal approximation

Stein's method is a collection of probabilistic techniques, using differential operators in order to assess quantities of the type

$$|E[h(F)] - E[h(Z)]|,$$

where  $Z$  and  $F$  are generic random variables, and the function  $h$  is such that the two expectations are well defined. In the specific case where  $Z \sim \mathcal{N}(0, 1)$ , with  $\mathcal{N}(0, 1)$  a standard Gaussian law, one is led to consider the so-called *Stein equation* associated with  $h$ , which is classically given by

$$h(x) - E[h(Z)] = f'(x) - xf(x), \quad x \in \mathbb{R}. \quad (2.23)$$

A solution to (2.23) is a function  $f$ , depending on  $h$ , which is Lebesgue a.e.-differentiable, and such that there exists a version of  $f'$  verifying (2.23) for every  $x \in \mathbb{R}$ . The following result collects findings by Stein [43, 44], Barbour [2], Daly [9] and Götze [15]. Precisely, the proof of Point (i) (which is known as *Stein's lemma*) involves a standard use of the Fubini theorem (see e.g. [8, Lemma 2.1]). Point (ii) is proved e.g. in [44, Lemma II.3] (for the estimates on the first and second derivative), in [9, Theorem 1.1] (for the estimate on the third derivative) and in [2] and [15] (for the alternative estimate on the first derivative). From here onwards, denote by  $\mathcal{C}_b^k$  the set of all real-valued bounded functions with bounded derivatives up to  $k^{\text{th}}$  order.

**Lemma 2.15.** (i) *Let  $W$  be a random variable. Then,  $W \stackrel{\text{Law}}{=} Z \sim \mathcal{N}(0, 1)$  if and only if for every continuous, piecewise continuously differentiable function  $f$  such that  $\mathbb{E}|f'(Z)| < \infty$ ,*

$$\mathbb{E}[f'(W) - Wf(W)] = 0. \quad (2.24)$$

(ii) If  $h \in \mathcal{C}_b^2$ , then (2.23) has a solution  $f_h$  which is thrice differentiable and such that  $\|f_h'\|_\infty \leq 4\|h\|_\infty$ ,  $\|f_h''\|_\infty \leq 2\|h'\|_\infty$  and  $\|f_h'''\|_\infty \leq 2\|h''\|_\infty$ . We also have  $\|f_h'\| \leq \|h''\|_\infty$ .

Now fix a function  $h \in \mathcal{C}_b^2$ , consider a Gaussian random variable  $Z \sim \mathcal{N}(0, 1)$ , and let  $F$  be a generic square integrable random variable. Integrating both sides of (2.23) with respect to the law of  $F$  gives

$$|E[h(F)] - E[h(Z)]| = |E[f_h'(F) - Ff_h(F)]| \quad (2.25)$$

with  $f_h$  the function given in Lemma 2.15. In the following sections we will show that, if  $F$  is a functional of the infinite Rademacher sequence  $X$ , then the quantity in the RHS of (2.25) can be successfully assessed by means of discrete Malliavin operators.

**Remark 2.16.** Plainly, if the sequence  $F_n$ ,  $n \geq 1$ , is such that  $|E[h(F_n)] - E[h(Z)]| \rightarrow 0$  for every  $h \in \mathcal{C}_b^2$ , then  $F_n \xrightarrow{\text{Law}} Z$ .

### 3 General Bounds for Rademacher functionals

#### 3.1 Main bound

The following result combines Proposition 2.14 with Lemma 2.15-(ii) and Relation (2.25) in order to estimate expressions of the type  $|E[h(F)] - E[h(Z)]|$ , where  $F$  is a square integrable functional of the infinite Rademacher sequence  $X$ , and  $Z \sim \mathcal{N}(0, 1)$ .

**Theorem 3.1.** Let  $F \in \text{dom}D$  be centered and such that  $\sum_k E|D_k F|^4 < \infty$ . Consider a function  $h \in \mathcal{C}_b^2$  and let  $Z \sim \mathcal{N}(0, 1)$ . Then,

$$|E[h(F)] - E[h(Z)]| \leq \min(4\|h\|_\infty, \|h''\|_\infty)B_1 + \|h''\|_\infty B_2, \quad (3.26)$$

where

$$\begin{aligned} B_1 &= E|1 - \langle DF, -DL^{-1}F \rangle_{\ell^2(\mathbb{N})}| \leq \sqrt{E[(1 - \langle DF, -DL^{-1}F \rangle_{\ell^2(\mathbb{N})})^2]}; \\ B_2 &= \frac{20}{3}E\left[\langle |DL^{-1}F|, |DF|^3 \rangle_{\ell^2(\mathbb{N})}\right]. \end{aligned} \quad (3.27)$$

*Proof.* Since  $h \in \mathcal{C}_b^2$ , Equality (2.25) holds. Observe that, since the first two derivatives of  $f_h$  are bounded, one has that  $f_h(F) \in \text{dom}D$  (this fact can be proved by using a Taylor expansion, as well as the assumptions on  $DF$  and the content of the first point of Remark 2.11). Using Lemma 2.12, one deduces that

$$E[f_h'(F) - Ff_h(F)] = E[f_h'(F) - \langle Df_h(F), -DL^{-1}F \rangle_{\ell^2(\mathbb{N})}].$$

We now use again the fact that  $f_h''$  is bounded: as an application of the first point of Lemma 2.13, we deduce therefore that

$$E[D_k L^{-1}F \times (f_h''(F_k^+) + f_h''(F_k^-))(D_k F)^2 X_k] = 0, \quad k \geq 1;$$

in particular, the boundedness of  $f_h''$ , along with the fact that  $E|D_k F|^4 < \infty$  and Lemma 2.13-(3), ensure that the previous expectation is well-defined. Finally, the desired conclusion follows from the chain rule proved in Proposition 2.14, as well as the bounds on  $\|f_h'\|_\infty$  and  $\|f_h'''\|_\infty$  stated in Lemma 2.15-(ii).  $\square$

**Remark 3.2.** 1. The requirement that  $\sum_k E|D_k F|^4 < \infty$  is automatically fulfilled whenever  $F$  belongs to a *finite* sum of chaoses. This can be deduced from the hypercontractive inequalities stated e.g. in [21, Ch. 6].

2. Since we are considering random variables possibly depending on the whole sequence  $X$  and having an infinite chaotic expansion, the expectation in (3.27) may actually be infinite.

3. Fix  $q \geq 1$  and let  $F$  have the form of a multiple integral of the type  $F = J_q(f)$ , where  $f \in \ell_0^2(\mathbb{N})^{\circ q}$ . Then, one has that

$$\langle DF, -DL^{-1}F \rangle_{\ell^2(\mathbb{N})} = \frac{1}{q} \|DF\|_{\ell^2(\mathbb{N})}^2, \quad (3.28)$$

$$\langle |DL^{-1}F|, |DF|^3 \rangle_{\ell^2(\mathbb{N})} = \frac{1}{q} \|DF\|_{\ell^4(\mathbb{N})}^4. \quad (3.29)$$

4. Let  $G$  be an isonormal Gaussian process (see [19] or [29]) over some separable Hilbert space  $\mathfrak{H}$ , and assume that  $F \in L^2(\sigma\{G\})$  is centered and differentiable in the sense of Malliavin. Then the Malliavin derivative of  $F$ , noted  $DF$ , is a  $\mathfrak{H}$ -valued random element, and in [25] the following upper bound is established (via Stein's method) between the laws of  $F$  and  $Z \sim \mathcal{N}(0, 1)$ :

$$d_{TV}(F, Z) \leq 2E|1 - \langle DF, -DL^{-1}F \rangle_{\mathfrak{H}}|,$$

where  $L^{-1}$  is here the inverse of the Ornstein-Uhlenbeck generator associated with  $G$  and  $d_{TV}$  is the total variation distance between the laws of  $F$  and  $Z$ .

5. Let  $N$  be a compensated Poisson measure over some measurable space  $(A, \mathcal{A})$ , with  $\sigma$ -finite control measure given by  $\mu$ . Assume that  $F \in L^2(\sigma\{N\})$  is centered and differentiable in the sense of Malliavin. Then the derivative of  $F$  is a random processes which a.s. belongs to  $L^2(\mu)$ . In [31, Section 3] the following upper bound for the Wasserstein distance (see Section 3.4) between the laws of  $F$  and  $Z \sim \mathcal{N}(0, 1)$  is proved (again by means of Stein's method):

$$d_W(F, Z) \leq E|1 - \langle DF, -DL^{-1}F \rangle_{L^2(\mu)}| + E \int_A |D_a F|^2 \times |D_a L^{-1}F| \mu(da).$$

6. In [6], Theorem 2.2, a bound to the normal distribution is given for functions  $F$  which strictly depend on the first  $n$  coordinates of  $X$  only. The bound is in terms of the operator  $\Delta_j F$  of  $F$ , see (2.17). Using the connection in Remark 2.11 point 2, the term corresponding to  $B_2$  in (3.27) would correspond to a bound on  $\frac{1}{n} \sum_{j=1}^n E|D_j F|$ . Instead of  $B_1$  in (3.27), [6] gives a bound in terms of the variance of a conditional expectation which depends on an exchangeable pair of sequences. Such variances are typically not easy to bound, except in the case of quadratic forms; see also Section 4.2.

### 3.2 First examples: Rademacher averages

A (possibly infinite) *Rademacher average* is just an element of the first chaos associated with  $X$ , that is, a random variable of the type

$$F = \sum_{i=1}^{\infty} \alpha_i X_i, \quad \text{with } \alpha \in \ell^2(\mathbb{N}). \quad (3.30)$$

See e.g. [22, Ch. IV] for general facts concerning random variables of the type (3.30). The next result is a consequence of Theorem 3.1 and of the relations (3.28)–(3.29). It yields a simple and explicit upper bound for the Gaussian approximation of Rademacher averages. We note that this simple example already goes beyond the typical examples in Stein’s method, as it treats a function  $F$  which depends on infinitely many coordinates of  $X$ .

**Corollary 3.3.** *Let  $F$  have the form (3.30), let  $Z \sim \mathcal{N}(0, 1)$  and consider  $h \in \mathcal{C}_b^2$ . Then*

$$|E[h(F)] - E[h(Z)]| \leq \min(4\|h\|_\infty, \|h''\|_\infty) \left| 1 - \sum_{i=1}^{\infty} \alpha_i^2 \right| + \frac{20}{3} \|h''\|_\infty \sum_{i=1}^{\infty} \alpha_i^4. \quad (3.31)$$

The proof of Corollary 3.3 (whose details are left to the reader) is easily deduced from the fact that, for  $F$  as in (3.30), one has that (3.28)–(3.29) with  $q = 1$  hold, and  $D_i F = \alpha_i$  for all  $i \geq 1$ .

As an example, for the special case of Corollary 3.3, if  $F_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n X_i$ , then it is easy to see that Relation (3.31) yields a bound of order  $n^{-1}$ , implying a faster rate than in the classical Berry-Esséen estimates. This faster rate arises from our use of smooth test functions; a related result is obtained in [17] using a coupling approach.

**Remark 3.4.** (A Mehler-type representation) Take an independent copy of  $X$ , noted  $X^* = \{X_1^*, X_2^*, \dots\}$ , fix  $t > 0$ , and consider the sequence  $X^t = \{X_1^t, X_2^t, \dots\}$  defined as follows:  $X^t$  is a sequence of i.i.d. random variables such that, for every  $k \geq 1$ ,  $X_k^t = X_k$  with probability  $e^{-t}$  and  $X_k^t = X_k^*$  with probability  $1 - e^{-t}$  (we stress that the choice between  $X$  and  $X^*$  is made separately for every  $k$ , so that one can have for instance  $X_1^t = X_1, X_2^t = X_2^*, \dots$  and so on). Then  $X^t$  is a Rademacher sequence and one has the following representation: for every  $F \in L^2(\sigma(X))$  and every  $\omega = (\omega_1, \omega_2, \dots) \in \Omega = \{-1, +1\}^{\mathbb{N}}$

$$P_t F(\omega) = E[F(X^t) | X = \omega], \quad (3.32)$$

where  $P_t$  is the Ornstein-Uhlenbeck semigroup given in (2.22). To prove such a representation of  $P_t$ , just consider a random variable of the type  $X_{j_1} \times \dots \times X_{j_d}$ , and then use a density argument. Now observe that, for a centered  $F = \sum_{n \geq 1} J_n(f_n) \in L_0^2(\sigma\{X\})$ , Equation (2.21) holds, and consequently

$$-D_k L^{-1} F = \sum_{n \geq 1} J_{n-1}(f_n(k, \cdot)) = \int_0^\infty e^{-t} P_t D_k F dt = \int_0^\infty e^{-t} E[D_k F(X^t) | X] dt,$$

so that

$$\langle DF, -DL^{-1} F \rangle_{\ell^2(\mathbb{N})} = \int_0^\infty e^{-t} \sum_{k \geq 1} D_k F E[D_k F(X^t) | X] dt = \int_0^\infty e^{-t} \langle DF, E[DF(X^t) | X] \rangle_{\ell^2(\mathbb{N})} dt.$$

Note that the representation (3.32) of the Ornstein-Uhlenbeck semigroup is not specific of Rademacher sequences, and could be e.g. extended to the case where  $X$  is i.i.d. standard Gaussian (see for instance [24, Section 2.2]). This construction would be different from the one leading to the usual *Mehler formula* (see e.g. [29, Formula (1.67)]). See Chatterjee [7, Lemma 1.1] and Nourdin and Peccati [25, Remark 3.6] for further connections between Mehler formulae and Stein’s method in a Gaussian setting.



### 3.3 A general exchangeable pair construction

Remark 3.4 uses a particular example of exchangeable pair of infinite Rademacher sequences. We shall now show that the chaos decomposition approach links in very well with the method of *exchangeable pairs*, and we also provide alternative bounds to [6] in terms of an exchangeable pair. Using a chaos decomposition, instead of a univariate approach such as suggested in [6], we employ a multivariate setting which conditions on exchangeable versions of the elements in the chaos decomposition.

Let us assume that  $F = F(X_1, \dots, X_d)$  is a random variable that depends uniquely on the first  $d$  coordinates  $\mathbf{X}_d = (X_1, \dots, X_d)$  of the Rademacher sequence  $X$ , with finite chaotic decomposition of the form (2.12). Now assume that  $E(F) = 0$  and  $E(F^2) = 1$  so that

$$F = \sum_{n=1}^d \sum_{1 \leq i_1 < \dots < i_n \leq d} n! f_n(i_1, \dots, i_n) X_{i_1} \cdots X_{i_n} = \sum_{n=1}^d J_n(f_n). \quad (3.33)$$

A natural exchangeable pair construction is as follows. Pick an index  $I$  at random, so that  $P(I = i) = \frac{1}{d}$  for  $i = 1, \dots, d$ , independently of  $X_1, \dots, X_d$ , and if  $I = i$  replace  $X_i$  by an independent copy  $X_i^*$  in all sums in the decomposition (3.33) which involve  $X_i$ . Call the resulting expression  $F'$ . Also denote the vector of Rademacher variables with the exchanged component by  $\mathbf{X}'_d$ . Then  $(F, F')$  forms an exchangeable pair.

In [37] an embedding approach is introduced, which suggests enhancing the pair  $(F, F')$  to a pair of vectors  $(\mathbf{W}, \mathbf{W}')$  which then ideally satisfy the linearity condition

$$E(\mathbf{W}' - \mathbf{W} | \mathbf{W}) = -\Lambda \mathbf{W} \quad (3.34)$$

with  $\Lambda$  being a deterministic matrix. In our example, choosing as embedding vector  $\mathbf{W} = (J_1(f_1), \dots, J_d(f_d))$ , we check that

$$\begin{aligned} & E(J'_n(f_n) - J_n(f_n) | \mathbf{W}) \\ &= -\frac{1}{d} \sum_{i=1}^d \sum_{1 \leq i_1 < \dots < i_n \leq d} \mathbf{1}_{\{i_1, \dots, i_n\}}(i) n! f_n(i_1, \dots, i_n) E(X_{i_1} \cdots X_{i_n} | \mathbf{W}) \\ &= -\frac{n}{d} J_n(f_n). \end{aligned}$$

Thus, with  $\mathbf{W}' = (J'_1(f_1), \dots, J'_d(f_d))$ , the condition (3.34) is satisfied, with  $\Lambda = (\lambda_{i,j})_{1 \leq i,j \leq d}$  being zero off the diagonal and  $\lambda_{n,n} = \frac{n}{d}$  for  $n = 1, \dots, d$ . Our random variable of interest  $F$  is a linear combination of the elements in  $\mathbf{W}$ , and we obtain the simple expression

$$E(F' - F | \mathbf{W}) = \frac{1}{d} L F = -\frac{1}{d} \delta D F. \quad (3.35)$$

This coupling helps to assess the distance to the normal distribution, as follows.

**Theorem 3.5.** *Let  $F \in \text{dom} D$  be centered and such that  $E(F^2) = 1$ . Let  $h \in \mathcal{C}_b^1$ , let  $Z \sim \mathcal{N}(0, 1)$ , and let  $(F, F')$  form the exchangeable pair constructed as above, which satisfies (3.35). Denote by  $L'$*

the Ornstein-Uhlenbeck operator for the exchanged Rademacher sequence  $\mathbf{X}'_d$ , and denote by  $(L')^{-1}$  its inverse. Then,

$$\begin{aligned} |E[h(F)] - E[h(Z)]| &\leq 4\|h\|_\infty \sqrt{\text{Var} \left[ \frac{d}{2} E \left[ (F' - F) \times ((L')^{-1}F' - L^{-1}F) \mid \mathbf{W} \right] \right]} \\ &\quad + \frac{d}{2} \|h'\|_\infty E \left[ (F' - F)^2 \times |(L')^{-1}F' - L^{-1}F| \right]. \end{aligned}$$

*Proof.* Let  $h \in \mathcal{C}_b^1$ , and let  $g$  denote the solution of the Stein equation (2.23) for  $h$ . Using antisymmetry, we have that for all smooth  $g$ ,

$$\begin{aligned} 0 &= E \left[ (g(F) + g(F')) \times ((L')^{-1}F' - L^{-1}F) \right] \\ &= 2E \left[ g(F)((L')^{-1}F' - L^{-1}F) \right] + E \left[ (g(F') - g(F)) \times ((L')^{-1}F' - L^{-1}F) \right] \\ &= \frac{2}{d} E[Fg(F)] + E \left[ (g(F') - g(F)) \times ((L')^{-1}F' - L^{-1}F) \right]; \end{aligned} \quad (3.36)$$

the last equality coming from (3.35) as well as the definitions of  $(L')^{-1}$  and  $L^{-1}$ . Hence

$$E[g'(F) - Fg(F)] = E \left[ g'(F) \left( 1 + \frac{d}{2} E \left[ (F' - F) \times ((L')^{-1}F' - L^{-1}F) \mid \mathbf{W} \right] \right) \right] + R,$$

where by Taylor expansion

$$|R| \leq \frac{d}{4} \|g''\|_\infty E \left[ (F' - F)^2 \times |(L')^{-1}F' - L^{-1}F| \right].$$

From (3.36) with  $g(x) = x$  we obtain

$$E \left[ E \left[ (F' - F) \times ((L')^{-1}F' - L^{-1}F) \mid \mathbf{W} \right] \right] = E \left[ (F' - F) \times ((L')^{-1}F' - L^{-1}F) \right] = -\frac{2}{d}$$

by our assumption that  $EF^2 = 1$ . We now only need to apply the Cauchy-Schwarz inequality and use the bounds from Lemma 2.15 to complete the proof.  $\square$

Equation (3.34) is key to answering the question on how to construct an exchangeable pair in the case of a random variable which admits a finite Rademacher chaos decomposition. Indeed, one can proceed as follows. Firstly, use  $\mathbf{W} = (J_1(f_1), \dots, J_d(f_d))$ , where the components come from the chaos decomposition. Then the results from [37] can be applied to assess the distance of  $\mathbf{W}$  to a normal distribution with the same covariance matrix. Note that due to the orthogonality of the integrals, the covariance matrix will be zero off the diagonal.

Finally, note that formally, using (3.35) and that  $LF = \lim_{t \rightarrow 0} \frac{P_t F - F}{t}$ ,

$$E(F(\mathbf{X}'_d) - F(\mathbf{X}_d) \mid \mathbf{W}) = \frac{1}{d} \lim_{t \rightarrow 0} \frac{P_t F(\mathbf{X}_d) - F(\mathbf{X}_d)}{t} = \frac{1}{d} \lim_{t \rightarrow 0} \frac{1}{t} [E(F(\mathbf{X}_d^t) \mid \mathbf{X}_d) - F(\mathbf{X}_d)],$$

where the notation is the same as in Remark 3.4. In this sense, our exchangeable pair can be viewed as the limit as  $t \rightarrow 0$  of the construction in Remark 3.4.

### 3.4 From twice differentiable functions to the Wasserstein distance

Although many of our results are phrased in terms of smooth test functions, we can also obtain results in Wasserstein distance by mimicking the smoothing in [38] used for Kolmogorov distance, but now apply it to Lipschitz functions. Recall that the Wasserstein distance between the laws of  $Y$  and  $Z$  is given by

$$d_W(Y, Z) = \sup_{h \in \text{Lip}(1)} |E[h(Y)] - E[h(Z)]|,$$

where  $\text{Lip}(1)$  is the class of all Lipschitz-continuous functions with Lipschitz constant less or equal to 1. Rademacher's Theorem states that a function which is Lipschitz continuous on the real line is Lebesgue-almost everywhere differentiable. For any  $h \in \text{Lip}(1)$  we denote by  $h'$  its derivative, which exists almost everywhere.

**Corollary 3.6.** *Let  $Z \sim \mathcal{N}(0, 1)$  and let  $F \in \text{dom}D$  be centered. Suppose that (3.26) holds for every function  $h \in \mathcal{C}_b^2$  and that  $4(B_1 + B_2) \leq 5$ . Then*

$$d_W(F, Z) \leq \sqrt{2(B_1 + B_2)(5 + E|F|)}.$$

*Proof.* Let  $h \in \text{Lip}(1)$  and, for  $t > 0$ , define

$$h_t(x) = \int_{-\infty}^{\infty} h(\sqrt{t}y + \sqrt{1-t}x)\phi(y)dy,$$

where  $\phi$  denotes the standard normal density. Then we may differentiate and integrate by parts, using that  $\phi'(x) = -x\phi(x)$ , to get

$$h_t''(x) = \frac{1-t}{\sqrt{t}} \int_{-\infty}^{\infty} yh'(\sqrt{t}y + \sqrt{1-t}x)\phi(y)dy.$$

Hence for  $0 < t < 1$  we may bound

$$\|h_t''\|_{\infty} \leq \frac{1-t}{\sqrt{t}} \|h'\|_{\infty} \int_{-\infty}^{\infty} |y|\phi(y)dy \leq \frac{1}{\sqrt{t}}. \quad (3.37)$$

Taylor expansion gives that for  $0 < t \leq \frac{1}{2}$  so that  $\sqrt{t} \leq \sqrt{1-t}$ ,

$$\begin{aligned} |E[h(F)] - E[h_t(F)]| &\leq \left| E \int \left\{ h(\sqrt{t}y + \sqrt{1-t}F) - h(\sqrt{1-t}F) \right\} \phi(y)dy \right| \\ &\quad + E \left| h(\sqrt{1-t}F) - h(F) \right| \\ &\leq \|h'\|_{\infty} \sqrt{t} \int |y|\phi(y)dy + \|h'\|_{\infty} \frac{t}{2\sqrt{1-t}} E|F| \leq \sqrt{t} \left\{ 1 + \frac{1}{2}E|F| \right\}. \end{aligned}$$

Here we used that  $\|h'\|_{\infty} \leq 1$  and that for  $0 < \theta < 1$ , we have  $(\sqrt{1-\theta t})^{-1} < (\sqrt{1-t})^{-1}$ . Similarly,  $|Eh(Z) - Eh_t(Z)| \leq \frac{3}{2}\sqrt{t}$ . Using (3.26) with (3.37) together with the triangle inequality we have for all  $h \in \text{Lip}(1)$

$$|E[h(F)] - E[h(Z)]| \leq \frac{1}{\sqrt{t}}(B_1(1) + B_2(1)) + \frac{1}{2}\sqrt{t}(5 + E|F|).$$

Minimising in  $t$  gives that the optimal is achieved for  $t = 2(B_1(1) + B_2(1))/(5 + E|F|)$ , and using this  $t$  yields the assertion.  $\square$

## 4 Normal approximation on a fixed chaos

### 4.1 Explicit upper bounds and CLTs

We now focus on random variables of the type  $F = J_q(f)$ , where  $J_q(f)$  is the multiple integral defined in Formula (2.6),  $q \geq 2$  and  $f \in \ell_0^2(\mathbb{N})^{\otimes q}$ . Due to the definition of the derivative operator  $D$ , as given in Section 2.5, we know that  $F \in \text{dom}D$ . Moreover, by combining Theorem 3.1 with formulae (3.28)–(3.29) one infers that, for every  $h \in \mathcal{C}_b^2$  and for  $Z \sim \mathcal{N}(0, 1)$ ,

$$\begin{aligned} & |E[h(F)] - E[h(Z)]| \\ & \leq \min(4\|h\|_\infty, \|h''\|_\infty) \sqrt{E \left[ \left( 1 - \frac{1}{q} \|DF\|_{\ell^2(\mathbb{N})}^2 \right)^2 \right]} + \frac{20}{3q} \|h''\|_\infty E\|DF\|_{\ell^4(\mathbb{N})}^4. \end{aligned} \quad (4.38)$$

The following statement provides an explicit bound of the two expectations appearing in (4.38).

**Theorem 4.1.** Fix  $q \geq 2$ , let  $f \in \ell_0^2(\mathbb{N})^{\otimes q}$ , and set  $F = J_q(f)$ . Then

$$\begin{aligned} & E \left\{ \left( 1 - \frac{1}{q} \|DF\|_{\ell^2(\mathbb{N})}^2 \right)^2 \right\} \\ & = \left| 1 - q! \|f\|_{\ell^2(\mathbb{N})^{\otimes q}}^2 \right|^2 \\ & \quad + q^2 \sum_{p=1}^{q-1} \left\{ (p-1)! \binom{q-1}{p-1}^2 \right\}^2 (2q-2p)! \| \widetilde{f \star_p^p f} \times \mathbf{1}_{\Delta_{2(q-p)}} \|_{\ell^2(\mathbb{N})^{\otimes 2(q-p)}}^2 \end{aligned} \quad (4.39)$$

$$\begin{aligned} & \leq \left| 1 - q! \|f\|_{\ell^2(\mathbb{N})^{\otimes q}}^2 \right|^2 \\ & \quad + q^2 \sum_{p=1}^{q-1} \left\{ (p-1)! \binom{q-1}{p-1}^2 \right\}^2 (2q-2p)! \|f \star_p^p f \times \mathbf{1}_{\Delta_{2(q-p)}}\|_{\ell^2(\mathbb{N})^{\otimes 2(q-p)}}^2, \end{aligned} \quad (4.40)$$

and

$$E\|DF\|_{\ell^4(\mathbb{N})}^4 \leq q^4 \sum_{p=1}^q \left\{ (p-1)! \binom{q-1}{p-1}^2 \right\}^2 (2q-2p)! \|f \star_p^{p-1} f\|_{\ell^2(\mathbb{N})^{\otimes (2(q-p)+1)}}^2. \quad (4.41)$$

**Remark 4.2.** According to Lemma 2.4, one has that

$$\|f \star_1^0 f\|_{\ell^2(\mathbb{N})^{\otimes (2q-1)}}^2 = \|f \star_q^{q-1} f\|_{\ell^2(\mathbb{N})}^2 \leq \|f\|_{\ell^2(\mathbb{N})^{\otimes q}}^2 \times \max_j \sum_{b_1, \dots, b_{q-1}} f^2(j, b_1, \dots, b_{q-1}),$$

and also, by combining this relation with (2.4), for  $p = 2, \dots, n-1$ ,

$$\|f \star_p^{p-1} f\|_{\ell^2(\mathbb{N})^{\otimes (2(q-p)+1)}}^2 \leq \|f\|_{\ell^2(\mathbb{N})^{\otimes q}}^3 \sqrt{\max_j \sum_{b_1, \dots, b_{q-1}} f^2(j, b_1, \dots, b_{q-1})}.$$

These estimates imply that, once  $\|f\|_{\ell^2(\mathbb{N})^{\otimes q}}$  is given, the bound on the RHS of (4.41) can be assessed by means of a uniform control on the “influence indices”  $\sum_{b_1, \dots, b_{q-1}} f^2(j, b_1, \dots, b_{q-1})$ ,  $j \geq 1$ . As

shown in [24, Theorem 2.1 and Section 3], when  $F = J_q(f)$  depends uniquely on a finite number of components of  $X$ , these influence indices can be used to directly measure the distance between the law of  $F$  and the law of the random variable, say  $F_G$ , obtained by replacing the Rademacher sequence with a i.i.d. Gaussian one. This result roughly implies that the two components of the bound (4.38) have a different nature, namely: (4.41) controls the distance between  $F$  and its Gaussian-based counterpart  $F_G$ , whereas (4.39)-(4.40) assess the distance between  $F_G$  and a standard Gaussian random variable.

*Proof of Theorem 4.1.* Observe that, by a classical approximation argument and by virtue of Lemma 2.6, it is sufficient to consider kernels  $f$  with support in a set of the type  $\{1, \dots, N\}^q$ , where  $N < \infty$  (this is for convenience only, since it allows to freely apply some Fubini arguments). Since  $D_j F = qJ_{q-1}(f(j, \cdot))$ , one has that (due to the multiplication formula (2.11))

$$\begin{aligned} (D_j F)^2 &= q^2 \sum_{r=0}^{q-1} r! \binom{q-1}{r}^2 J_{2(q-1-r)} \left( f(j, \cdot) \star_r^{\widetilde{f}} f(j, \cdot) \times \mathbf{1}_{\Delta_{2(q-1-r)}} \right) \\ &= q^2 \sum_{p=1}^q (p-1)! \binom{q-1}{p-1}^2 J_{2(q-p)} \left( f(j, \cdot) \star_{p-1}^{\widetilde{f}} f(j, \cdot) \times \mathbf{1}_{\Delta_{2(q-p)}} \right). \end{aligned} \quad (4.42)$$

It follows that, by a Fubini argument,

$$\begin{aligned} \frac{1}{q} \|DF\|_{\ell^2(\mathbb{N})}^2 &= q \sum_{p=1}^q (p-1)! \binom{q-1}{p-1}^2 J_{2(q-p)} \left( f \star_{p-1}^{\widetilde{f}} f \times \mathbf{1}_{\Delta_{2(q-p)}} \right) \\ &= q! \|f\|_{\ell^2(\mathbb{N})^{\otimes q}}^2 + q \sum_{p=1}^{q-1} (p-1)! \binom{q-1}{p-1}^2 J_{2(q-p)} \left( f \star_{p-1}^{\widetilde{f}} f \times \mathbf{1}_{\Delta_{2(q-p)}} \right), \end{aligned}$$

and the equality (4.39) is obtained by means of orthogonality and isometric properties of multiple integrals. Inequality (4.40) is a consequence of the fact that, for any kernel  $h$ ,  $\|\widetilde{h}\| \leq \|h\|$ . To prove (4.41), use again (4.42) in order to write

$$\begin{aligned} &E[(D_j F)^4] \\ &= q^4 \sum_{p=1}^q \left\{ (p-1)! \binom{q-1}{p-1}^2 \right\}^2 (2q-2p)! \left\| f(j, \cdot) \star_{p-1}^{\widetilde{f}} f(j, \cdot) \times \mathbf{1}_{\Delta_{2(q-p)}} \right\|_{\ell^2(\mathbb{N})^{\otimes 2(q-p)}}^2 \\ &\leq q^4 \sum_{p=1}^q \left\{ (p-1)! \binom{q-1}{p-1}^2 \right\}^2 (2q-2p)! \left\| f(j, \cdot) \star_{p-1}^{\widetilde{f}} f(j, \cdot) \right\|_{\ell^2(\mathbb{N})^{\otimes 2(q-p)}}^2. \end{aligned}$$

The conclusion is obtained by using the identity

$$\sum_{j \in \mathbb{N}} \left\| f(j, \cdot) \star_{p-1}^{\widetilde{f}} f(j, \cdot) \right\|_{\ell^2(\mathbb{N})^{\otimes 2(q-p)}}^2 = \left\| f \star_{p-1}^{\widetilde{f}} f \right\|_{\ell^2(\mathbb{N})^{\otimes (2(q-p)+1)}}^2.$$

□

We deduce the following result, yielding sufficient conditions for CLTs on a fixed chaos.

**Proposition 4.3.** Fix  $q \geq 2$ , and let  $F_k = J_q(f_k)$ ,  $k \geq 1$ , be a sequence of multiple integrals such that  $f_k \in \ell_0^2(\mathbb{N})^{\otimes q}$  and  $E(F_k^2) = q! \|f_k\|_{\ell^2(\mathbb{N})^{\otimes q}} \rightarrow 1$ . Then, a sufficient condition in order to have that

$$F_k \xrightarrow{\text{Law}} Z \sim \mathcal{N}(0, 1) \tag{4.43}$$

is that

$$\|f_k \star_r^r f_k\|_{\ell^2(\mathbb{N})^{\otimes 2(q-r)}} \rightarrow 0, \quad \forall r = 1, \dots, q - 1. \tag{4.44}$$

*Proof.* One only needs to combine Theorem 4.1 and Relation (4.38) with Point 3 of Lemma 2.4.  $\square$

**Remark 4.4.** The content of Proposition 4.3 echoes the results proved in [30], where it is shown that, on a fixed Gaussian chaos of order  $q$ , the convergence to zero of the contractions  $\|f_k \star_r^r f_k\|$ ,  $r = 1, \dots, q - 1$ , is necessary and sufficient in order to have the CLT (4.43). See also [25] for bounds, involving norms of contractions, on the normal approximation of the elements of a Gaussian chaos. See [31] and [32] for analogous results concerning the normal approximation of regular functionals of a Poisson measure. Finally, observe that we do not know at the present time whether the condition (4.44) is necessary in order to have the CLT (4.43).

## 4.2 More on finite quadratic forms

When specialized to the case of normalized double integrals, Theorem 3.1, Proposition 4.3 and Remark 2.5 yield the following result.

**Corollary 4.5.** Let  $F_k = J_2(f_k)$ ,  $k \geq 1$ , be a sequence of double integrals such that  $E(F_k^2) = 2 \|f_k\|_{\ell^2(\mathbb{N})^{\otimes 2}}^2 = 1$  (for simplicity). Let  $Z \sim \mathcal{N}(0, 1)$ . Then, for every  $k \geq 1$  and every  $h \in \mathcal{C}_b^2$  it holds that

$$|E[h(F_k)] - E[h(Z)]| \tag{4.45}$$

$$\leq 4\sqrt{2} \min(4\|h\|_\infty, \|h''\|_\infty) \times \|f_k \star_1^1 f_k \times \mathbf{1}_{\Delta_2}\|_{\ell^2(\mathbb{N})^{\otimes 2}} + 160\|h''\|_\infty \times \|f_k \star_2^2 f_k\|_{\ell^2(\mathbb{N})}^2$$

$$\leq 4\sqrt{2} \min(4\|h\|_\infty, \|h''\|_\infty) \times \|f_k \star_1^1 f_k \times \mathbf{1}_{\Delta_2}\|_{\ell^2(\mathbb{N})^{\otimes 2}}$$

$$+ 160\|h''\|_\infty \times \|f_k \star_1^1 f_k \times \mathbf{1}_{\Delta_2^c}\|_{\ell^2(\mathbb{N})^{\otimes 2}}^2, \tag{4.46}$$

and a sufficient condition in order to have the CLT (4.43) is that  $\|f_k \star_1^1 f_k\|_{\ell^2(\mathbb{N})^{\otimes 2}} \rightarrow 0$ .

Now consider kernels  $f_k \in \ell_0^2(\mathbb{N})^{\otimes 2}$ ,  $k \geq 1$ , such that for every  $k$  the following holds: (i)  $2 \|f_k\|_{\ell^2(\mathbb{N})^{\otimes 2}}^2 = 1$ , and (ii) the support of  $f_k$  is contained in the set  $\{1, \dots, k\}^2$ . Then, the random variables

$$F_k = J_2(f_k) = \sum_{1 \leq i, j \leq k} f_k(i, j) X_i X_j, \quad k \geq 1, \tag{4.47}$$

are quadratic functionals (with no diagonal terms) of the vectors  $(X_1, \dots, X_k)$ . Limit theorems involving sequences such as (4.47), for general vectors of i.i.d. random variables, have been intensively studied in the probabilistic literature – see e.g. [10], [16] and the references therein. The following result, providing a complete characterization of CLTs for sequences such as (4.47), can be deduced from the main findings contained in [10] (but see also [11] for extensions to general multilinear functionals).

**Proposition 4.6.** *Let  $X$  be the Rademacher sequence considered in this paper, and let  $F_k$  be given by (4.47). For every  $k \geq 1$ , write  $[f_k]$  for the  $k \times k$  square symmetric matrix  $\{f_k(i, j) : 1 \leq i, j \leq k\}$ . Then, the following three conditions are equivalent as  $k \rightarrow \infty$ :*

- (a)  $F_k \xrightarrow{\text{Law}} Z \sim \mathcal{N}(0, 1)$ ;
- (b)  $\text{Trace}\{[f_k]^4\} = \|f_k \star_1^1 f_k\|_{\ell^2(\mathbb{N})^{\otimes 2}}^2 \rightarrow 0$ ;
- (c)  $E(F_k^4) \rightarrow E(Z^4) = 3$ .

**Remark 4.7.** Condition (b) in the previous statement is often replaced by the following:

$$(b') \text{Trace}\{[f_k]^4\} \rightarrow 0 \quad \text{and} \quad \max_{j \leq k} \sum_{i=1}^k f_k(i, j)^2 \rightarrow 0.$$

However, Relation (2.5) shows that the second requirement in (b') is indeed redundant.

In [6], Chatterjee proves that, in Wasserstein distance,

$$d_W(F_k, Z) \leq \sqrt{\frac{1}{2} \text{Trace}\{[f_k]^4\}} + \frac{5}{2} \sum_{j=1}^k \left[ \sum_{i=1}^k f_k(i, j)^2 \right]^{\frac{3}{2}}. \quad (4.48)$$

Note that

$$\max_{j \leq k} \left[ \sum_{i=1}^k f_k(i, j)^2 \right]^{\frac{3}{2}} \leq \sum_{j=1}^k \left[ \sum_{i=1}^k f_k(i, j)^2 \right]^{\frac{3}{2}} \leq \frac{1}{2} \max_{j \leq k} \left[ \sum_{i=1}^k f_k(i, j)^2 \right]^{\frac{1}{2}}, \quad (4.49)$$

and due e.g. to (2.5), Relation (4.48) gives an alternate proof of the implication (b)  $\rightarrow$  (a) in Proposition 4.6. Another proof of the same implication can be deduced from the following result, which is a direct consequence of Corollary 4.5 and Lemma 2.4 (Point 3).

**Corollary 4.8.** *Let the notation and assumptions of this section prevail. Then, for every  $k$  and for every  $h \in \mathcal{C}_b^2$ , one has that*

$$\begin{aligned} & |E[h(F_k)] - E[h(Z)]| \\ & \leq 4\sqrt{2} \min(4\|h\|_\infty, \|h''\|_\infty) \|f_k \star_1^1 f_k \times \mathbf{1}_{\Delta_2}\|_{\ell^2(\mathbb{N})^{\otimes 2}} + 160\|h''\|_\infty \|f_k \star_1^1 f_k \times \mathbf{1}_{\Delta_2^c}\|_{\ell^2(\mathbb{N})^{\otimes 2}}^2 \\ & \leq 4\sqrt{2} \min(4\|h\|_\infty, \|h''\|_\infty) \sqrt{\text{Trace}\{[f_k]^4\}} + 160\|h''\|_\infty \text{Trace}\{[f_k]^4\}. \end{aligned}$$

## 5 Sums of single and double integrals and applications to runs

### 5.1 Motivation: infinite weighted runs

Due to their simple dependence structure, runs, and more generally,  $m$ -dependent variables, lend themselves to normal approximations; for applications see for example [1]. The first Berry-Esséen bound for overlapping success runs was derived in [14]. In the context of Stein's method, the sum of  $m$ -dependent variables is one of the original examples from Stein's 1972 paper [43]; in this paper, it was treated by a local approach. Ever since, runs can be viewed as a benchmark example for Stein's

method. Some instances of its treatment include [39], where an exchangeable pair coupling is employed for a univariate normal approximation, in the wider context of weighted U-statistics, and [37], where exchangeable pairs are used for a multivariate normal approximation, in both cases for the count runs of finite length. The paper [8] uses a local approach in connection with smoothing to treat  $m$ -dependent random fields. Typically these studies focus on counting runs in a finite window of a possibly infinite sequence, see for example [1] for details. In contrast, our method allows to consider infinite weighted sums of runs.

Let  $\xi = \{\xi_n : n \in \mathbb{Z}\}$  be a standard *Bernoulli sequence*. This means that  $\xi$  is a collection of i.i.d. random variables such that  $P[\xi_1 = 1] = P[\xi_1 = 0] = 1/2$ . In the sequel, we will suppose without loss of generality that  $\xi = \frac{1}{2}(1 + X)$  where  $X = \{X_n : n \in \mathbb{Z}\}$  is a two-sided *Rademacher* sequence, obtained e.g. by juxtaposing two independent Rademacher sequences of the kind introduced in Section 2.1. Note that we chose  $\mathbb{Z}$  as a parameter set, in order to simplify the forthcoming discussion. Also, in what follows we will implicitly use the fact that all the results and bounds established in the previous sections extend to this setting, by simply replacing sums over  $\mathbb{N}$  with sums over  $\mathbb{Z}$  (the proof of this elementary fact is left to the reader). Now fix an integer  $d \geq 1$  as well as a sequence  $\{\alpha^{(n)} : n \geq 1\}$  of elements of  $\ell^2(\mathbb{Z})$ . In this section, we study the normal approximation of the sequence  $\{G_n : n \geq 1\}$  defined by

$$G_n = \sum_{i \in \mathbb{Z}} \alpha_i^{(n)} \xi_i \dots \xi_{i+d}. \quad (5.50)$$

We call  $G_n$  an *infinite weighted  $d$ -run*. Actually, we will rather focus on the normalized version of  $G_n$ , namely

$$F_n = \frac{G_n - E(G_n)}{\sqrt{\text{Var}G_n}}. \quad (5.51)$$

Observe that, in order to apply Theorem 3.1, one needs to find the chaotic expansion of  $F_n$  or, equivalently, that of  $G_n$ . Using the identity  $\xi = \frac{1}{2}(1 + X)$ , it is immediately deduced, see also Section 2.4-(ii), that

$$\begin{aligned} G_n &= 2^{-(d+1)} \sum_{i \in \mathbb{Z}} \alpha_i^{(n)} (1 + X_i) \dots (1 + X_{i+d}) \\ &= 2^{-(d+1)} \sum_{i \in \mathbb{Z}} \alpha_i^{(n)} \sum_{I \subset \{i, \dots, i+d\}} X_{i_1} \dots X_{i_{|I|}} \quad (\text{with } I = \{i_1, \dots, i_{|I|}\}) \\ &= 2^{-(d+1)} \sum_{r=0}^{d+1} \sum_{i \in \mathbb{Z}} \alpha_i^{(n)} \sum_{\substack{I \subset \{i, \dots, i+d\} \\ |I|=r}} X_{i_1} \dots X_{i_r} \\ &= E(G_n) + \sum_{r=1}^{d+1} J_r \left( 2^{-(d+1)} \sum_{i \in \mathbb{Z}} \alpha_i^{(n)} \sum_{\substack{I \subset \{i, \dots, i+d\} \\ |I|=r}} \mathbf{1}_{\{i_1\}} \widetilde{\otimes} \dots \widetilde{\otimes} \mathbf{1}_{\{i_r\}} \right), \end{aligned} \quad (5.52)$$

where the tilde indicates a symmetrization. In particular, it is now immediate to compute  $DG_n$  and  $DL^{-1}G_n$ , by using the definitions given in Section 2.5. However, since the analysis of (5.50) is meant to be only an illustration, from now on we will focus on the case where  $d = 1$  (2-runs). The general case could be handled in a similar way (at the cost of a quite cumbersome notation).



## 5.2 Normal approximation of sums of single and double integrals

We will deduce a bound for the quantity  $|E[h(F_n)] - E[h(Z)]|$  (where  $Z \sim \mathcal{N}(0, 1)$  and  $F_n$  is given by (5.51)) from the following result, which can be seen as a particular case of Theorem 3.1.

**Proposition 5.1.** (Sum of a single and a double integral) *Let  $F = J_1(f) + J_2(g)$  with  $f \in \ell^2(\mathbb{Z})$  and  $g \in \ell_0^2(\mathbb{Z})^{\otimes 2}$ . Assume that  $\sum_{i \in \mathbb{Z}} |g(i, k)| < \infty$  for all  $k \in \mathbb{Z}$ . Also, suppose (for simplicity) that  $\text{Var}F = 1$ , and let  $h \in \mathcal{C}_b^2$ . Then*

$$\begin{aligned} |E[h(F)] - E[h(Z)]| &\leq \min(4\|h\|_\infty, \|h''\|_\infty) (2\sqrt{2}\|g \star_1^1 g \mathbf{1}_{\Delta_2}\|_{\ell^2(\mathbb{Z})^{\otimes 2}} + 3\|f \star_1^1 g\|_{\ell^2(\mathbb{Z})}) \\ &\quad + \frac{160}{3}\|h''\|_\infty \sum_{k \in \mathbb{Z}} \left[ f^4(k) + 16 \left( \sum_{i \in \mathbb{Z}} |g(i, k)| \right)^4 \right]. \end{aligned} \quad (5.53)$$

**Remark 5.2.** By applying successively Fubini and Cauchy-Schwarz theorems, one has that

$$\|f \star_1^1 g\|_{\ell^2(\mathbb{Z})}^2 = \sum_{i, j \in \mathbb{Z}} f(i)f(j) g \star_1^1 g(i, j) \leq \|f\|_{\ell^2(\mathbb{Z})}^2 \|g \star_1^1 g\|_{\ell^2(\mathbb{Z})^{\otimes 2}}. \quad (5.54)$$

This inequality has an interesting consequence. Suppose indeed that the sequences  $J_1(f_n)$  and  $J_2(g_n)$ ,  $n \geq 1$ , are such that, as  $n \rightarrow \infty$ : (a)  $\|f_n\|_{\ell^2(\mathbb{Z})} \rightarrow 1$ , (b)  $2\|g_n\|_{\ell^2(\mathbb{Z})^{\otimes 2}}^2 \rightarrow 1$ , (c)  $\sum_{k \in \mathbb{Z}} f_n^4(k) \rightarrow 0$ , (d)  $\sum_{k \in \mathbb{Z}} \left( \sum_{i \in \mathbb{Z}} |g_n(i, k)| \right)^4 \rightarrow 0$ , and (e)  $\|g_n \star_1^1 g_n\|_{\ell^2(\mathbb{Z})^{\otimes 2}} \rightarrow 0$ . Then the estimate (5.54) and Proposition 5.1 imply that, for every  $(\alpha, \beta) \neq (0, 0)$ , the sequence

$$\frac{\alpha J_1(f_n) + \beta J_2(g_n)}{\sqrt{\alpha^2 + \beta^2}}, \quad n \geq 1,$$

converges in law to  $Z \sim \mathcal{N}(0, 1)$ , and therefore that the vectors  $(J_1(f_n), J_2(g_n))$ ,  $n \geq 1$ , jointly converge in law towards a two-dimensional centered i.i.d. Gaussian vector with unit variances. Note that each one of conditions (a)–(e) involves *separately* one of the two kernels  $f_n$  and  $g_n$ . See [27] and [31, Section 6], respectively, for several related results in a Gaussian and in a Poisson framework.

*Proof of Proposition 5.1.* Firstly, observe that  $D_k F = 2J_1(g(\cdot, k)) + f(k) = 2 \sum_{i \in \mathbb{Z}} g(i, k) X_i + f(k)$  so that, using  $(a + b)^4 \leq 8(a^4 + b^4)$ ,

$$E|D_k F|^4 \leq 128E \left( \sum_{i \in \mathbb{Z}} g(i, k) X_i \right)^4 + 8f^4(k) \leq 128 \left( \sum_{i \in \mathbb{Z}} |g(i, k)| \right)^4 + 8f^4(k) < \infty.$$

We are thus left to bound  $B_1$  and  $B_2$  in Theorem 3.1, taking into account the particular form of  $F$ . We have  $-L^{-1}F = \frac{1}{2}J_2(g) + J_1(f)$  so that  $-D_k L^{-1}F = J_1(g(\cdot, k)) + f(k)$ . Consequently, with the multiplication formula (2.11), we get

$$\begin{aligned} &\langle DF, -DL^{-1}F \rangle_{\ell^2(\mathbb{Z})} \\ &= 2 \sum_{k \in \mathbb{Z}} J_1(g(\cdot, k))^2 + 3 \sum_{k \in \mathbb{Z}} f(k) J_1(g(\cdot, k)) + \|f\|_{\ell^2(\mathbb{Z})}^2 \\ &= 2 \sum_{k \in \mathbb{Z}} J_2(g(\cdot, k) \otimes g(\cdot, k) \mathbf{1}_{\Delta_2}) + 3 \sum_{k \in \mathbb{Z}} f(k) J_1(g(\cdot, k)) + \|f\|_{\ell^2(\mathbb{Z})}^2 + 2\|g\|_{\ell^2(\mathbb{Z})^{\otimes 2}}^2 \\ &= 2J_2(g \star_1^1 g \mathbf{1}_{\Delta_2}) + 3J_1(f \star_1^1 g) + \|f\|_{\ell^2(\mathbb{Z})}^2 + 2\|g\|_{\ell^2(\mathbb{Z})^{\otimes 2}}^2 \end{aligned}$$

so that

$$\begin{aligned} E\left|\langle DF, -DL^{-1}F \rangle_{\ell^2(\mathbb{Z})} - \text{Var}F\right|^2 &= E\left|2J_2(g \star_1^1 g \mathbf{1}_{\Delta_2}) + 3J_1(f \star_1^1 g)\right|^2 \\ &= 8\|g \star_1^1 g \mathbf{1}_{\Delta_2}\|_{\ell^2(\mathbb{Z})^{\otimes 2}}^2 + 9\|f \star_1^1 g\|_{\ell^2(\mathbb{Z})}^2. \end{aligned}$$

Hence,

$$\begin{aligned} B_1 &\leq \sqrt{8\|g \star_1^1 g \mathbf{1}_{\Delta_2}\|_{\ell^2(\mathbb{Z})^{\otimes 2}}^2 + 9\|f \star_1^1 g\|_{\ell^2(\mathbb{Z})}^2} \\ &\leq (2\sqrt{2}\|g \star_1^1 g \mathbf{1}_{\Delta_2}\|_{\ell^2(\mathbb{Z})^{\otimes 2}} + 3\|f \star_1^1 g\|_{\ell^2(\mathbb{Z})}). \end{aligned}$$

Now, let us consider  $B_2$ . We have

$$|D_k F| \leq |f(k)| + 2 \sum_{i \in \mathbb{Z}} |g(i, k)|.$$

Similarly,

$$\begin{aligned} |D_k L^{-1}F| &= \left|f(k) + J_1(g(\cdot, k))\right| = \left|f(k) + \sum_{i \in \mathbb{Z}} g(i, k)X_i\right| \\ &\leq |f(k)| + \sum_{i \in \mathbb{Z}} |g(i, k)| \leq |f(k)| + 2 \sum_{i \in \mathbb{Z}} |g(i, k)|. \end{aligned}$$

Still using  $(a + b)^4 \leq 8(a^4 + b^4)$ , we deduce

$$\begin{aligned} \sum_{k \in \mathbb{Z}} |D_k L^{-1}F| \times |D_k F|^3 &\leq \sum_{k \in \mathbb{Z}} \left(|f(k)| + 2 \sum_{i \in \mathbb{Z}} |g(i, k)|\right)^4 \\ &\leq 8 \sum_{k \in \mathbb{Z}} \left[f^4(k) + 16 \left(\sum_{i \in \mathbb{Z}} |g(i, k)|\right)^4\right]. \end{aligned}$$

Hence

$$B_2 \leq \frac{160}{3} \sum_{k \in \mathbb{Z}} \left[f^4(k) + 16 \left(\sum_{i \in \mathbb{Z}} |g(i, k)|\right)^4\right]$$

and the desired conclusion follows by applying Theorem 3.1.  $\square$

### 5.3 Bounds for infinite 2-runs

When  $d = 1$ , Proposition 5.1 allows to deduce the following bound for the normal approximation of the random variable  $F_n$  defined in (5.51).

**Proposition 5.3.** *Let  $\{F_n : n \geq 1\}$  be the sequence defined by  $F_n = \frac{G_n - E(G_n)}{\sqrt{\text{Var}G_n}}$  with*

$$G_n = \sum_{i \in \mathbb{Z}} \alpha_i^{(n)} \xi_i \xi_{i+1}.$$

Here,  $\xi = \{\xi_n : n \in \mathbb{Z}\}$  stands for the standard Bernoulli sequence and  $\{\alpha^{(n)} : n \geq 1\}$  is a given sequence of elements of  $\ell^2(\mathbb{Z})$ . Consider a function  $h \in \mathcal{C}_b^2$ . Then, for  $Z \sim \mathcal{N}(0, 1)$ ,

$$\begin{aligned} |E[h(F)] - E[h(Z)]| &\leq \frac{7}{16} \times \frac{\min(4\|h\|_\infty, \|h''\|_\infty)}{\text{Var}G_n} \times \sqrt{\sum_{i \in \mathbb{Z}} (\alpha_i^{(n)})^4} \\ &\quad + \frac{35}{24} \times \frac{\|h''\|_\infty}{(\text{Var}G_n)^2} \times \sum_{i \in \mathbb{Z}} (\alpha_i^{(n)})^4 \end{aligned} \quad (5.55)$$

with

$$\text{Var}G_n = \frac{3}{16} \sum_{i \in \mathbb{Z}} (\alpha_i^{(n)})^2 + \frac{1}{8} \sum_{i \in \mathbb{Z}} \alpha_i^{(n)} \alpha_{i+1}^{(n)}. \quad (5.56)$$

It follows that a sufficient condition to have  $F_n \xrightarrow{\text{Law}} Z$  is that

$$\sum_{i \in \mathbb{Z}} (\alpha_i^{(n)})^4 = o\left((\text{Var}G_n)^2\right) \quad \text{as } n \rightarrow \infty.$$

*Proof.* Identity (5.56) is easily verified. On the other hand, by (5.52), we have

$$F_n = \frac{G_n - E(G_n)}{\sqrt{\text{Var}G_n}} = J_1(f) + J_2(g),$$

with

$$\begin{aligned} f &= \frac{1}{4\sqrt{\text{Var}G_n}} \sum_{a \in \mathbb{Z}} \alpha_a^{(n)} (\mathbf{1}_{\{a\}} + \mathbf{1}_{\{a+1\}}) \\ g &= \frac{1}{8\sqrt{\text{Var}G_n}} \sum_{a \in \mathbb{Z}} \alpha_a^{(n)} (\mathbf{1}_{\{a\}} \otimes \mathbf{1}_{\{a+1\}} + \mathbf{1}_{\{a+1\}} \otimes \mathbf{1}_{\{a\}}). \end{aligned}$$

Now, let us compute each quantity appearing in the RHS of (5.53). If  $i \neq j$  then

$$(g \star_1^1 g)(i, j) = \frac{1}{64\text{Var}G_n} \left( \alpha_i^{(n)} \alpha_{i+1}^{(n)} \mathbf{1}_{\{j=i+2\}} + \alpha_j^{(n)} \alpha_{j+1}^{(n)} \mathbf{1}_{\{j=i-2\}} \right).$$

Hence

$$\begin{aligned} &\|g \star_1^1 g \mathbf{1}_{\Delta_2}\|_{\ell^2(\mathbb{Z})^{\otimes 2}} \\ &= \frac{1}{64\text{Var}G_n} \sqrt{\sum_{i, j \in \mathbb{Z}} \left[ (\alpha_i^{(n)})^2 (\alpha_{i+1}^{(n)})^2 \mathbf{1}_{\{j=i+2\}} + (\alpha_j^{(n)})^2 (\alpha_{j+1}^{(n)})^2 \mathbf{1}_{\{j=i-2\}} \right]} \\ &= \frac{\sqrt{2}}{64\text{Var}G_n} \sqrt{\sum_{i \in \mathbb{Z}} (\alpha_i^{(n)})^2 (\alpha_{i+1}^{(n)})^2}. \end{aligned}$$

We have

$$(f \star_1^1 g)(i) = \frac{1}{32\text{Var}G_n} \left( \alpha_i^{(n)} \alpha_{i+1}^{(n)} + (\alpha_{i-1}^{(n)})^2 + (\alpha_i^{(n)})^2 + \alpha_{i-1}^{(n)} \alpha_{i-2}^{(n)} \right).$$

Hence, using  $(a + b + c + d)^2 \leq 4(a^2 + b^2 + c^2 + d^2)$ ,

$$\begin{aligned} & \|f \star_1^1 g\|_{\ell^2(\mathbb{Z})} \\ & \leq \frac{1}{16\text{Var}G_n} \sqrt{\sum_{i \in \mathbb{Z}} (\alpha_i^{(n)})^2 (\alpha_{i+1}^{(n)})^2 + \sum_{i \in \mathbb{Z}} (\alpha_{i-1}^{(n)})^4 + \sum_{i \in \mathbb{Z}} (\alpha_i^{(n)})^4 + \sum_{i \in \mathbb{Z}} (\alpha_{i-1}^{(n)})^2 (\alpha_{i-2}^{(n)})^2} \\ & = \frac{\sqrt{2}}{16\text{Var}G_n} \sqrt{\sum_{i \in \mathbb{Z}} (\alpha_i^{(n)})^2 (\alpha_{i+1}^{(n)})^2 + \sum_{i \in \mathbb{Z}} (\alpha_i^{(n)})^4}. \end{aligned}$$

We have, using  $(a + b)^4 \leq 8(a^4 + b^4)$ ,

$$\begin{aligned} \sum_{k \in \mathbb{Z}} f^4(k) &= \frac{1}{256(\text{Var}G_n)^2} \sum_{k \in \mathbb{Z}} \left[ \sum_{a \in \mathbb{Z}} \alpha_a^{(n)} (\mathbf{1}_{\{a\}}(k) + \mathbf{1}_{\{a+1\}}(k)) \right]^4 \\ &\leq \frac{1}{16(\text{Var}G_n)^2} \sum_{k \in \mathbb{Z}} (\alpha_k^{(n)})^4. \end{aligned}$$

Finally, still using  $(a + b)^4 \leq 8(a^4 + b^4)$ ,

$$\begin{aligned} & \sum_{k \in \mathbb{Z}} \left[ \sum_{i \in \mathbb{Z}} |g(i, k)| \right]^4 \\ & \leq \frac{1}{4096(\text{Var}G_n)^2} \sum_{k \in \mathbb{Z}} \left[ \sum_{i, a \in \mathbb{Z}} |\alpha_a^{(n)}| |\mathbf{1}_{\{a\}}(i) \mathbf{1}_{\{a+1\}}(k) + \mathbf{1}_{\{a+1\}}(i) \mathbf{1}_{\{a\}}(k)| \right]^4 \\ & \leq \frac{1}{256(\text{Var}G_n)^2} \sum_{k \in \mathbb{Z}} (\alpha_k^{(n)})^4. \end{aligned}$$

Now, the desired conclusion follows by plugging all these estimates in (5.53), after observing that  $\sum_{i \in \mathbb{Z}} (\alpha_i^{(n)})^2 (\alpha_{i+1}^{(n)})^2 \leq \sum_{i \in \mathbb{Z}} (\alpha_i^{(n)})^4$ , by the Cauchy-Schwarz inequality.  $\square$

## 6 Multiple integrals over sparse sets

### 6.1 General results

Fix  $d \geq 2$ . Let  $F_N$ ,  $N \geq 1$ , be a sequence of subsets of  $\mathbb{N}^d$  such that the following three properties are satisfied for every  $N \geq 1$ : (i)  $F_N \neq \emptyset$ , (ii)  $F_N \subset \Delta_d^N$  (as defined in (2.2)), that is,  $F_N$  is contained in  $\{1, \dots, N\}^d$  and has no diagonal components, and (iii)  $F_N$  is a symmetric set, in the sense that every  $(i_1, \dots, i_d) \in F_N$  is such that  $(i_{\sigma(1)}, \dots, i_{\sigma(d)}) \in F_N$  for every permutation  $\sigma$  of the set  $\{1, \dots, d\}$ . Let  $X$  be the infinite Rademacher sequence considered in this paper. Given sets  $F_N$  as at points (i)–(iii), we shall consider the sequence of multilinear forms

$$\tilde{S}_N = [d! \times |F_N|]^{-\frac{1}{2}} \sum_{(i_1, \dots, i_d) \in F_N} X_{i_1} \cdots X_{i_d} = J_d(f_N), \quad N \geq 1, \quad (6.57)$$

where  $|F_N|$  stands for the cardinality of  $F_N$ , and

$$f_N(i_1, \dots, i_d) := [d! \times |F_N|]^{-\frac{1}{2}} \times \mathbf{1}_{F_N}(i_1, \dots, i_d).$$

Note that  $E(\tilde{S}_N) = 0$  and  $E(\tilde{S}_N^2) = 1$  for every  $N$ . In the paper [5], Blei and Janson studied the problem of finding conditions on the set  $F_N$ , in order to have that the CLT

$$\tilde{S}_N \xrightarrow{\text{Law}} Z \sim \mathcal{N}(0, 1), \quad N \rightarrow \infty, \quad (6.58)$$

holds.

**Remark 6.1.** Strictly speaking, Blei and Janson use the notation  $F_N$  in order to indicate the *restriction to the simplex*  $\{(i_1, \dots, i_d) : i_1 < i_2 < \dots < i_d\}$  of a set verifying Properties (i)–(iii) above.

In order to state Blei and Janson’s result, we need to introduce some more notation.

**Remark on notation.** In what follows, we will write  $\mathbf{a}_k$  to indicate vectors  $\mathbf{a}_k = (a_1, \dots, a_k)$  belonging to a set of the type  $\{1, \dots, N\}^k =: [N]^k$ , for some  $k, N \geq 1$ . We will regard these objects both as vectors and sets, for instance: an expression of the type  $\mathbf{a}_k \cap \mathbf{i}_l = \emptyset$ , means that the two sets  $\{a_1, \dots, a_k\}$  and  $\{i_1, \dots, i_l\}$  have no elements in common; when writing  $j \in \mathbf{a}_k$ , we mean that  $j = a_r$  for some  $r = 1, \dots, k$ ; when writing  $\mathbf{a}_k \subset \mathbf{i}_l$  ( $k \leq l$ ), we indicate that, for every  $r = 1, \dots, k$ , one has  $a_r = i_s$  for some  $s = 1, \dots, l$ . When a vector  $\mathbf{a}_k$  enters in a sum, we will avoid to specify  $\mathbf{a}_k \in [N]^k$ , whenever the domain of summation  $[N]^k$  is clear from the context.

Given  $N \geq 1$  and an index  $j \in [N]$ , we set

$$F_{N,j}^* = \{\mathbf{i}_d \in F_N : j \in \mathbf{i}_d\}.$$

For every  $N$ , the set  $F_N^\# \subset F_N \times F_N$  is defined as the collection of all pairs  $(\mathbf{i}_d, \mathbf{k}_d) \in F_N \times F_N$  such that: (a)  $\mathbf{i}_d \cap \mathbf{k}_d = \emptyset$ , and (b) there exists  $p = 1, \dots, d - 1$ , as well as  $\mathbf{i}'_p \subset \mathbf{i}_d$  and  $\mathbf{k}'_p \subset \mathbf{k}_d$  such that

$$((\mathbf{k}'_p, \mathbf{i}_d \setminus \mathbf{i}'_p), (\mathbf{i}'_p, \mathbf{k}_d \setminus \mathbf{k}'_p)) \in F_N \times F_N,$$

where  $\mathbf{i}_d \setminus \mathbf{i}'_p$  represents the element of  $[N]^{d-p}$  obtained by eliminating from  $\mathbf{i}_d$  the coordinates belonging to  $\mathbf{i}'_p$ , and  $(\mathbf{k}'_p, \mathbf{i}_d \setminus \mathbf{i}'_p)$  is the element of  $[N]^d$  obtained by replacing  $\mathbf{i}'_p$  with  $\mathbf{k}'_p$  in  $\mathbf{i}_d$  (an analogous description holds for  $(\mathbf{i}'_p, \mathbf{k}_d \setminus \mathbf{k}'_p)$ ). In other words, the  $2d$  indices  $i_1, \dots, i_d, j_1, \dots, j_d$  can be partitioned in at least two ways into elements of  $F_N$ .

**Theorem 6.2.** ([5, Th. 1.7]) *Let the above notation and assumptions prevail, and suppose that*

$$\lim_{N \rightarrow \infty} \max_{j \leq N} \frac{|F_{N,j}^*|}{|F_N|} = 0, \quad \text{and} \quad (6.59)$$

$$\lim_{N \rightarrow \infty} \frac{|F_N^\#|}{|F_N|^2} = 0. \quad (6.60)$$

*Then Relation (6.58) holds, with convergence of all moments.*

**Remark 6.3.** As pointed out in [5], Condition (6.60) can be described as a weak “sparseness condition” (see e.g. [4]). See also [5, Th. 1.7] for a converse statement.

The principal achievement of this section is the following refinement of Theorem 6.2.

**Theorem 6.4.** Under the above notation and assumptions, consider a function  $h \in \mathcal{C}_b^2$ . Then there exist two universal constants  $C_1$  and  $C_2$ , depending only on  $d$ ,  $\|h\|_\infty$  and  $\|h''\|_\infty$ , such that, for every  $N \geq 1$  and for  $Z \sim \mathcal{N}(0, 1)$ ,

$$|E[h(Z)] - E[h(\tilde{S}_N)]| \leq C_1 \frac{|F_N^\#|^{\frac{1}{2}}}{|F_N|} + C_2 \left[ \max_{j \leq N} \frac{|F_{N,j}^*|}{|F_N|} \right]^{\frac{1}{4}}. \quad (6.61)$$

*Proof.* By combining (4.38) with Theorem 4.1, we know that there exist universal combinatorial constants  $\alpha_p^d > 0$ ,  $p = 1, \dots, d-1$  and  $\beta_l^d > 0$ ,  $l = 1, \dots, d$ , such that

$$|E[h(\tilde{S}_N)] - E[h(Z)]| \leq \|h\|_\infty \sum_{p=1}^{d-1} \alpha_p^d \|f_N \star_p^p f_N \mathbf{1}_{\Delta_{2(d-p)}^N}\|_{\ell^2(\mathbb{N})^{\otimes 2(d-p)}} \quad (6.62)$$

$$+ \|h''\|_\infty \sum_{l=1}^{d-1} \beta_l^d \|f_N \star_l^{l-1} f_N\|_{\ell^2(\mathbb{N})^{\otimes 2(d-l)+1}}. \quad (6.63)$$

We now evaluate each norm appearing in (6.62)–(6.63). According to the second and the third points of Lemma 2.4 and using the fact that  $\|f_N\|_{\ell^2(\mathbb{N})^{\otimes d}}^2 = (d!)^{-1}$ , one has that

$$\begin{aligned} \|f_N \star_1^0 f_N\|_{\ell^2(\mathbb{N})^{\otimes (2d-1)}}^2 &= \|f_N \star_d^{d-1} f_N\|_{\ell^2(\mathbb{N})}^2 \\ &\leq \frac{1}{d!} \max_{j \leq N} \sum_{\mathbf{b}_{d-1}} f_N(j, \mathbf{b}_{d-1})^2 = \frac{1}{d \times d!} \max_{j \leq N} \frac{|F_{N,j}^*|}{|F_N|}. \end{aligned} \quad (6.64)$$

Also, by combining the two inequalities in the second point of Lemma 2.4, we deduce that, for every  $l = 2, \dots, d-1$ ,

$$\begin{aligned} \|f_N \star_l^{l-1} f_N\|_{\ell^2(\mathbb{N})^{\otimes 2(d-l)+1}}^2 &\leq \frac{1}{d!} \|f_N \star_d^{d-1} f_N\|_{\ell^2(\mathbb{N})}^2 \\ &\leq \frac{1}{(d!)^{3/2} \sqrt{d}} \left[ \max_{j \leq N} \frac{|F_{N,j}^*|}{|F_N|} \right]^{\frac{1}{2}} \leq \frac{1}{(d!)^{3/2} \sqrt{d}} \left[ \max_{j \leq N} \frac{|F_{N,j}^*|}{|F_N|} \right]^{\frac{1}{4}}. \end{aligned} \quad (6.65)$$

Now fix  $p = 1, \dots, d-1$ . One has that

$$\begin{aligned} &\|f_N \star_p^p f_N \mathbf{1}_{\Delta_{2(d-p)}^N}\|_{\ell^2(\mathbb{N})^{\otimes 2(d-p)}}^2 \\ &= \frac{1}{(d!|F_N|)^2} \sum_{\substack{\mathbf{a}_{d-p}, \mathbf{b}_p \\ \mathbf{x}_{d-p}, \mathbf{y}_p}} \mathbf{1}_{F_N}(\mathbf{a}_{d-p}, \mathbf{b}_p) \mathbf{1}_{F_N}(\mathbf{x}_{d-p}, \mathbf{b}_p) \mathbf{1}_{F_N}(\mathbf{x}_{d-p}, \mathbf{y}_p) \mathbf{1}_{F_N}(\mathbf{a}_{d-p}, \mathbf{y}_p) \mathbf{1}_{\{\mathbf{a}_{d-p} \cap \mathbf{x}_{d-p} = \emptyset\}} \\ &= \sum_{\gamma=0}^p \gamma! \binom{p}{\gamma}^2 \frac{U_\gamma}{(d!|F_N|)^2}, \end{aligned} \quad (6.66)$$

where

$$\begin{aligned} U_0 &= \sum_{\substack{\mathbf{a}_{d-p}, \mathbf{b}_p \\ \mathbf{x}_{d-p}, \mathbf{y}_p}} \mathbf{1}_{F_N}(\mathbf{a}_{d-p}, \mathbf{b}_p) \mathbf{1}_{F_N}(\mathbf{x}_{d-p}, \mathbf{b}_p) \mathbf{1}_{F_N}(\mathbf{x}_{d-p}, \mathbf{y}_p) \mathbf{1}_{F_N}(\mathbf{a}_{d-p}, \mathbf{y}_p) \times \\ &\quad \times \mathbf{1}_{\{\mathbf{a}_{d-p} \cap \mathbf{x}_{d-p} = \emptyset\}} \mathbf{1}_{\{\mathbf{b}_p \cap \mathbf{y}_p = \emptyset\}} \\ &\leq |F_N^\#|, \end{aligned}$$

and

$$\begin{aligned} U_p &= \sum_{\substack{\mathbf{a}_{d-p}, \mathbf{b}_p \\ \mathbf{x}_{d-p}}} \mathbf{1}_{F_N}(\mathbf{a}_{d-p}, \mathbf{b}_p) \mathbf{1}_{F_N}(\mathbf{x}_{d-p}, \mathbf{b}_p) \mathbf{1}_{\{\mathbf{a}_{d-p} \cap \mathbf{x}_{d-p} = \emptyset\}} \\ &\leq \sum_{\mathbf{b}_p} \left[ \sum_{\mathbf{a}_{d-p}} \mathbf{1}_{F_N}(\mathbf{a}_{d-p}, \mathbf{b}_p) \right]^2 \leq |F_N| \times \max_{j \leq N} |F_{N,j}^*|, \end{aligned}$$

and finally, for  $\gamma = 1, \dots, p-1$ ,

$$\begin{aligned} U_\gamma &= \sum_{\substack{\mathbf{a}_{d-p}, \mathbf{u}_\gamma, \mathbf{b}_{p-\gamma} \\ \mathbf{x}_{d-p}, \mathbf{y}_{p-\gamma}}} \mathbf{1}_{F_N}(\mathbf{a}_{d-p}, \mathbf{u}_\gamma, \mathbf{b}_{p-\gamma}) \mathbf{1}_{F_N}(\mathbf{x}_{d-p}, \mathbf{u}_\gamma, \mathbf{b}_{p-\gamma}) \times \\ &\quad \times \mathbf{1}_{F_N}(\mathbf{x}_{d-p}, \mathbf{u}_\gamma, \mathbf{y}_{p-\gamma}) \mathbf{1}_{F_N}(\mathbf{a}_{d-p}, \mathbf{u}_\gamma, \mathbf{y}_{p-\gamma}) \mathbf{1}_{\{\mathbf{a}_{d-p} \cap \mathbf{x}_{d-p} = \emptyset\}} \mathbf{1}_{\{\mathbf{b}_{p-\gamma} \cap \mathbf{y}_{p-\gamma} = \emptyset\}} \\ &\leq \sum_{\substack{\mathbf{a}_{d-p}, \mathbf{u}_\gamma, \mathbf{b}_{p-\gamma} \\ \mathbf{x}_{d-p}, \mathbf{y}_{p-\gamma}}} \mathbf{1}_{F_N}(\mathbf{a}_{d-p}, \mathbf{u}_\gamma, \mathbf{b}_{p-\gamma}) \mathbf{1}_{F_N}(\mathbf{x}_{d-p}, \mathbf{u}_\gamma, \mathbf{b}_{p-\gamma}) \times \\ &\quad \times \mathbf{1}_{F_N}(\mathbf{x}_{d-p}, \mathbf{u}_\gamma, \mathbf{y}_{p-\gamma}) \mathbf{1}_{F_N}(\mathbf{a}_{d-p}, \mathbf{u}_\gamma, \mathbf{y}_{p-\gamma}) \mathbf{1}_{\{\mathbf{b}_{p-\gamma} \cap \mathbf{y}_{p-\gamma} = \emptyset\}} \\ &\leq |F_N|^2 \times \|f_N \star_{d-p+1}^{d-p} f_N\|_{\ell^2(\mathbb{N})^{\otimes 2p-1}}^2. \end{aligned}$$

The proof is concluded by using the estimates (6.64) (for  $p = 1$ ) and (6.65) (for  $p = 2, \dots, d-1$ ).  $\square$

We shall now state a generalization of Theorem 6.4, providing an explicit upper bound for the normal approximation of multiple integrals defined over *infinite* sets. This result shows once again that our approach allows to deal directly with the Gaussian approximation of random variables that are functions of the whole *infinite* sequence  $X$ . Consider a sequence of real numbers  $\beta = \{\beta_i : i \geq 1\} \in \ell^2(\mathbb{N})$ , and define the finite measure on  $\mathbb{N}$

$$m_\beta(A) = \sum_{i \in A} \beta_i^2, \quad A \subset \mathbb{N}. \quad (6.67)$$

We denote by  $m_\beta^d$  ( $d \geq 2$ ) the canonical  $d$ -product measure associated with  $m_\beta$ . For every  $d \geq 2$  and for every set  $F \subset \mathbb{N}^d$ , we define the sets  $F^\# \subset F \times F$  and  $F_j^*$ ,  $j = 1, 2, \dots$  as before. In particular,  $F_j^*$  is the collection of all  $(i_1, \dots, i_d) \in F$  such that  $j = i_k$  for some  $k$ . For  $\beta$  as before, and  $F \in \Delta_d$  (possibly infinite) such that  $m_\beta(F) > 0$ , we are interested in the normal approximation of the random variable

$$J(\beta, F) = \frac{1}{[d!m_\beta(F)]^{1/2}} \sum_{(i_1, \dots, i_d) \in F} \beta_{i_1} \cdots \beta_{i_d} X_{i_1} \cdots X_{i_d}.$$

The following statement, whose proof (omitted) follows along the lines of that of Theorem 5.4, provides an upper bound for the normal approximation of  $J(\beta, F)$ .

**Proposition 6.5.** *Let  $Z \sim \mathcal{N}(0, 1)$ . Under the above notation and assumptions, for every  $h \in \mathcal{C}_b^2$  one has that there exist positive constants  $K_1$  and  $K_2$ , depending uniquely on  $d$ ,  $\|h\|_\infty$  and  $\|h'\|_\infty$ , such that*

$$\left| E[h(Z)] - E[h(J(\beta, F))] \right| \leq K_1 \frac{m_\beta^{2d}(F^\#)^{\frac{1}{2}}}{m_\beta^d(F)} + K_2 \left[ \sup_{j \geq 1} \frac{m_\beta^d(F_j^*)}{m_\beta^d(F)} \right]^{\frac{1}{4}}. \quad (6.68)$$

## 6.2 Fractional Cartesian products

In this section, we describe an explicit application of Theorem 6.4. The framework and notation are basically the same as those of Example 1.2 in [5]. Fix integers  $d \geq 3$  and  $2 \leq m \leq d-1$ , and consider a collection  $\{S_1, \dots, S_d\}$  of distinct non-empty subsets of  $[d] = \{1, \dots, d\}$  such that: (i)  $S_i \neq \emptyset$ , (ii)  $\bigcup_i S_i = [d]$ , (iii)  $|S_i| = m$  for every  $i$ , (iv) each index  $j \in [d]$  appears in exactly  $m$  of the sets  $S_i$ , and (v) the cover  $\{S_1, \dots, S_d\}$  is connected (i.e., it cannot be partitioned into two disjoint partial covers). For every  $i = 1, \dots, d$  and every  $\mathbf{y}_d = (y_1, \dots, y_d) \in \mathbb{N}^d$ , we use the notation  $\pi_{S_i} \mathbf{y} = (y_j : j \in S_i)$ . Note that the operator  $\pi_{S_i}$  transforms vectors of  $\mathbb{N}^d$  into vectors of  $\mathbb{N}^m$ . For every  $N \geq d^m$ , write  $n = \lfloor N^{1/m} \rfloor$ , that is,  $n$  is the largest integer such that  $n \leq N^{1/m}$ . Now select a one-to-one map  $\varphi$  from  $[n]^m$  into  $[N]$ , and define

$$F_N^* = \{(\varphi(\pi_{S_1} \mathbf{k}_d), \dots, \varphi(\pi_{S_d} \mathbf{k}_d)) : \mathbf{k}_d \in [n]^d\} \subset [N]^d,$$

(note that, in general,  $F_N^*$  is *not* a symmetric set),  $F_N^{**} = F_N^* \cap \Delta_N^d$ , and also

$$F_N = \mathbf{sym}(F_N^{**}), \quad (6.69)$$

where  $\mathbf{sym}(F_N^{**})$  indicates the collections of all vectors  $\mathbf{y}_d = (y_1, \dots, y_d) \in \mathbb{N}^d$  such that

$$(y_{\sigma(1)}, \dots, y_{\sigma(d)}) \in F_N^{**}$$

for some permutation  $\sigma$ . Then the sequence  $\{F_N\}$  is called a *fractional Cartesian product*. Fractional Cartesian products arose first in harmonic analysis, where they were shown to fill gaps between integer powers of spectral sets, see [4].

**Proposition 6.6.** *Let  $Z \sim \mathcal{N}(0, 1)$ , and let  $F_N$  and  $\tilde{S}_N$ ,  $N \geq d^m$ , be respectively defined according to (6.69) and (6.57). Then, for every  $h \in \mathcal{C}_b^2$ , there exists a constant  $K > 0$ , independent of  $N$ , such that*

$$|E[h(\tilde{S}_N)] - E[h(Z)]| \leq \frac{K}{N^{1/(2m)}}.$$

*Proof.* The computations contained in [5, p. 16] can be straightforwardly adapted to our setting, so that we deduce that the sequence  $\{F_N : N \geq d^m\}$  has *combinatorial dimension*  $\alpha = d/m$ . Recall that this means that there exist finite constants  $0 < Q_2 < Q_1 < \infty$  (independent of  $N$ ) such that: (a)  $|F_N| \geq Q_2 N^\alpha$ , and (b) for every  $A_1, \dots, A_d \subset [N]$ ,  $|F_N \cap (A_1 \times \dots \times A_d)| \leq Q_1 (\max_{1 \leq j \leq d} |A_j|)^\alpha$ . Thanks to Theorem 6.4, to conclude the proof it is therefore sufficient to check that, as  $N \rightarrow \infty$ ,

$$\max_{j \leq N} |F_{N,j}^*| = O(N^{\alpha-1}) \quad \text{and} \quad |F_N^\#| = O(N^{2\alpha-1/m}). \quad (6.70)$$

Start by observing that every element  $(z_1, \dots, z_d)$  of  $F_N$  has the form

$$(z_1, \dots, z_d) = (\varphi(\pi_{S_{\sigma(1)}} \mathbf{k}_d), \dots, \varphi(\pi_{S_{\sigma(d)}} \mathbf{k}_d)), \quad (6.71)$$

where  $\mathbf{k}_d = (k_1, \dots, k_d) \in [n]^d$  and  $\sigma$  is a permutation of  $[d]$ . Since  $\varphi$  is one-to-one, it follows that, for every  $j \leq N$ , there are at most  $d! d \times n^{d-m}$  elements of the set  $F_{N,j}$ . To see this, just observe that, every  $(z_1, \dots, z_d) \in F_{N,j}$  is completely specified by the following three elements: (i) a permutation  $\sigma$  of  $d$ , (ii) the index  $a = 1, \dots, d$  such that  $z_a = j$ , and (iii) the values of those coordinates  $k_b$  such that  $b \notin S_{\sigma(a)}$ . Since  $n \leq N^{1/m}$  by construction, one immediately obtains the first relation in (6.70).



To prove the second part of (6.70), we shall first show that, if  $((z_1, \dots, z_d), (z'_1, \dots, z'_d)) \in F_N^\#$  are such that  $(z_1, \dots, z_d)$  is as in (6.71), and  $(z'_1, \dots, z'_d) = (\varphi(\pi_{S_{\sigma'(1)}} \mathbf{k}'_d), \dots, \varphi(\pi_{S_{\sigma'(d)}} \mathbf{k}'_d))$  ( $\mathbf{k}'_d \in [n]^d$ ) then  $\mathbf{k}_d$  and  $\mathbf{k}'_d$  must have  $m$  coordinates in common. The definition of  $F_N^\#$  implies indeed that, for such a vector  $((z_1, \dots, z_d), (z'_1, \dots, z'_d))$ , there exists

$$(u_1, \dots, u_d) = (\varphi(\pi_{S_{\rho(1)}} \mathbf{i}_d), \dots, \varphi(\pi_{S_{\rho(d)}} \mathbf{i}_d)) \in F_N,$$

such that, for some  $p = 1, \dots, d - 1$ : (a) there exist indices  $a_1, \dots, a_p$  and  $b_1, \dots, b_p$  such that  $\pi_{S_{\rho(a_i)}} \mathbf{i}_d = \pi_{S_{\sigma(b_i)}} \mathbf{k}_d$  for every  $i = 1, \dots, p$ , and (b) there exist indices  $v_1, \dots, v_{d-p}$  and  $w_1, \dots, w_{d-p}$  such that  $\pi_{S_{\rho(v_i)}} \mathbf{i}_d = \pi_{S_{\sigma'(w_i)}} \mathbf{k}'_d$  for every  $i = 1, \dots, d - p$ . Note that then we necessarily have  $\{a_1, \dots, a_p, v_1, \dots, v_{d-p}\} = \{b_1, \dots, b_p, w_1, \dots, w_{d-p}\} = [d]$ . By connectedness, there exists at least one  $q \in [d]$  such that  $i_q$  (i.e., the  $q$ th coordinate of  $\mathbf{i}_d$ ) belongs both to one of the sets  $\pi_{S_{\rho(a_i)}} \mathbf{i}_d$  and to one of the sets  $\pi_{S_{\rho(v_i)}} \mathbf{i}_d$ . This last property implies that there exists a constant  $L$ , independent of  $N$ , such that  $|F_N^\#| \leq L n^{2d-1} \leq L N^{2\alpha-1/m}$ . This concludes the proof.  $\square$

**Remark 6.7.** Note that the combinatorial dimension  $\alpha = d/m$ , as appearing in the proof of Proposition 6.6, is an index of interdependence between the coordinates of the sets  $F_N$ . See Ch. XIII in [4] for more details on this point.

### 6.3 Beyond the Rademacher case: a question by Blei and Janson

Now we go back to the framework and notation of Section 6.1, so that, in particular, the sequence  $\tilde{S}_N$ ,  $N \geq 1$ , is defined according to (6.57). For every  $N$  define  $\tilde{S}_N^G$  to be the random variable obtained from (6.57) by replacing the sequence  $X$  with a i.i.d. Gaussian sequence  $G = \{G_i : i \geq 1\}$ , where each  $G_i$  has mean zero and unit variance. A natural question, which has been left open by Blei and Janson in [5, Remark 4.6], is whether under the conditions (6.59)–(6.60) the sequence  $\tilde{S}_N^G$ ,  $N \geq 1$ , converges in law towards a standard Gaussian distribution. Note that this problem could be tackled by a direct computation, based for instance on [25] or [30]. However, the results of this paper, combined with those of [24], allow to elegantly deduce a more general result, which we provide in the forthcoming statement. In what follows, we write  $V = \{V_i : i \geq 1\}$  to indicate a centered i.i.d. sequence, with unit variance and such that  $E|V_1|^3 = \eta < \infty$  (note that the results of [24] would allow to obtain similar results in even more general frameworks, but we do not look for generality here). We also denote by  $\tilde{S}_N^V$  the random variable obtained from (6.57) by replacing  $X$  with  $V$ .

**Proposition 6.8.** *Under the above notation and assumptions, consider a function  $h \in \mathcal{C}_b^3$ . Then, there exist two universal constants  $B_1$  and  $B_2$ , depending uniquely on  $d$ ,  $\eta$ ,  $\|h\|_\infty$ ,  $\|h''\|_\infty$  and  $\|h'''\|_\infty$ , such that, for every  $N \geq 1$  and for  $Z \sim \mathcal{N}(0, 1)$ ,*

$$|E[h(Z)] - E[h(\tilde{S}_N^V)]| \leq B_1 \frac{|F_N^\#|^{\frac{1}{2}}}{|F_N|} + B_2 \left[ \max_{j \leq N} \frac{|F_{Nj}^*|}{|F_N|} \right]^{\frac{1}{4}}. \quad (6.72)$$

*In particular, if (6.59)–(6.60) take place, then  $\tilde{S}_N^V$  converges in law towards  $Z$ .*

*Proof.* One has that

$$|E[h(Z)] - E[h(\tilde{S}_N^V)]| \leq |E[h(Z)] - E[h(\tilde{S}_N)]| + |E[h(\tilde{S}_N)] - E[h(\tilde{S}_N^V)]|,$$

and the conclusion is obtained by combining Theorem 6.4 with the fact that, according to [24, Theorem 3.18, case H2], there exists a constant  $Q$ , depending only on  $\|h'''\|_\infty$  and  $\eta$ , such that

$$\left| E[h(\tilde{S}_N)] - E[h(\tilde{S}_N^V)] \right| \leq Q \left[ \max_{j \leq N} \frac{|F_{Nj}^*|}{|F_N|} \right]^{\frac{1}{2}}.$$

□

**Acknowledgments.** This paper originates from all three authors attending the “Workshop on Stein’s method” at the Institute of Mathematical Sciences of the National University of Singapore, from March 31 until April 4, 2008. We warmly thank A. Barbour, L. Chen, K.-P. Choi and A. Xia for their kind hospitality and generous support. We are grateful to J. Dedecker for useful discussions. We would also like to thank two anonymous referees and the Associate Editor for helpful comments.

## References

- [1] N. Balakrishnan and M. V. Koutras (2002). *Runs and scans with applications*. Wiley Series in Probability and Statistics. Wiley-Interscience [John Wiley & Sons], New York. MR1882476 MR1882476
- [2] A.D. Barbour (1990). Stein’s method for diffusion approximations. *Probab. Theory Rel. Fields* **84**(3), 297-322. MR1035659 MR1035659
- [3] V. Bentkus, B.-Y. Jing and W. Zhou (2009). On normal approximations to U-statistics. *Ann. Probab.* **37**, 2174-2199. MR2573555 MR2573555
- [4] R. Blei (2001). *Analysis in Integer and Fractional Dimensions*. Cambridge University Press. MR1853423 MR1853423
- [5] R. Blei and S. Janson (2004). Rademacher chaos: tail estimates versus limit theorems. *Ark. Mat.* **42**(1), 13-29. MR2056543 MR2056543
- [6] S. Chatterjee (2008). A new method of normal approximation. *Ann Probab.* **36**, 1584-1610. MR2435859 MR2435859
- [7] S. Chatterjee (2009). Fluctuations of eigenvalues and second order Poincaré inequalities. *Probab. Theory Rel. Fields* **143**, 1-40. MR2449121 MR2449121
- [8] L.H.Y. Chen and Q.-M. Shao (2005). Stein’s method for normal approximation. In: *An introduction to Stein’s method*, 1-59. Lect. Notes Ser. Inst. Math. Sci. Natl. Univ. Singap. **4**, Singapore Univ. Press, Singapore, 2005. MR2235448 MR2235448
- [9] F. Daly (2008). Upper bounds for Stein-type operators. *Electron. J. Probab.* **13**, 566-587 (electronic). MR2399291 MR2399291
- [10] P. de Jong (1987). A central limit theorem for generalized quadratic forms. *Probab. Theory Rel. Fields* **75**(2), 261-277. MR0885466 MR0885466

- [11] P de Jong (1990). A central limit theorem for generalized multilinear forms. *J. Mult. Anal.* **34**, 275-289. MR1073110 MR1073110
- [12] V.H. de la Peña and E. Giné (1997). *Decoupling*. Springer-Verlag, Berlin Heidelberg New York. MR1666908
- [13] B. Efron and C. Stein (1981). The Jackknife Estimate of Variance. *Ann. Statist.* **9**, 586-596. MR0615434 (82k:62074) MR0615434
- [14] A.P Godbole (1992). The exact and asymptotic distribution of overlapping success runs. *Comm. Statist. - Theory and Methods* **21**, 953–967. MR1173302 MR1173302
- [15] F Götze (1991). On the rate of convergence in the multivariate CLT. *Ann. Probab.* **19**, 724-739. MR1106283 MR1106283
- [16] F Götze and A.N. Tikhomirov (2002). Asymptotic distribution of quadratic forms and applications. *J. Theoret. Probab.* **15**(2), 423-475. MR1898815 MR1898815
- [17] L. Goldstein and G. Reinert (1997). Stein’s method and the zero bias transformation with application to simple random sampling. *Ann. Appl. Probab.* **7**(4), 935-952. MR1484792 MR1484792
- [18] J. Hájek (1968). Asymptotic Normality of Simple Linear Rank Statistics Under Alternatives *Ann. Math. Statist.* **39**, 325-346. MR0222988 MR0222988
- [19] S. Janson (1997). *Gaussian Hilbert Spaces*. Cambridge University Press. MR1474726 MR1474726
- [20] S. Karlin and Y. Rinott (1982). Applications of ANOVA type decompositions for comparisons of conditional variance statistics including jackknife estimates. *Ann. Statist.* **10** 485-501. MR0653524 MR0653524
- [21] S. Kwapień and W.A. Woyczyński (1992). *Random Series and Stochastic Integrals: Single and Multiple*. Birkhäuser, Basel. MR1167198 MR1167198
- [22] M. Ledoux and M. Talagrand (1991). *Probability on Banach spaces*. Springer-Verlag, Berlin Heidelberg New York. MR1102015 MR1102015
- [23] P-A. Meyer (1992). *Quantum probability for probabilists*. LNM **1538**. Springer-Verlag, Berlin Heidelberg New York. MR1222649 MR1222649
- [24] E. Mossel, R. O’Donnell and K. Oleszkiewicz (2010). Noise stability of functions with low influences: variance and optimality. *Ann. Math.* **71**, 295–341. MR2630040 MR2630040
- [25] I. Nourdin and G. Peccati (2009). Stein’s method on Wiener chaos. *Probab. Theory Rel. Fields* **145**, no. 1, 75–118. MR2520122 MR2520122
- [26] I. Nourdin and G. Peccati (2010). Stein’s method and exact Berry-Esséen bounds for functionals of Gaussian fields. *Ann. Probab.* **37**, 2231–2261. MR2573557

- [27] I. Nourdin, G. Peccati and A. Réveillac (2008). Multivariate normal approximation using Stein's method and Malliavin calculus. *Ann. Inst. H. Poincaré Probab. Statist.* **46** (2010), 45–58. MR2641769
- [28] I. Nourdin and F.G. Viens (2009). Density estimates and concentration inequalities with Malliavin calculus. *Electron. J. Probab.* **14**, 2287–2309. MR2556018 MR2556018
- [29] D. Nualart (2006). *The Malliavin calculus and related topics*. Springer-Verlag, Berlin, 2nd edition. MR2200233 MR2200233
- [30] D. Nualart and G. Peccati (2005). Central limit theorems for sequences of multiple stochastic integrals. *Ann. Probab.* **33** (1), 177-193. MR2118863 MR2118863
- [31] G. Peccati, J.-L. Solé, M.S. Taqqu and F. Utzet (2010). Stein's method and normal approximation of Poisson functionals. *Ann. Probab.* **38**, 443-478. MR2642882 MR2642882
- [32] G. Peccati and M.S. Taqqu (2008). Central limit theorems for double integrals. *Bernoulli* **14**(3), 791-821. MR2537812 MR2537812
- [33] G. Peccati and M.S. Taqqu (2010). Wiener chaos: moments, cumulants and diagrams. Springer-Verlag, to appear.
- [34] N. Privault (2008). Stochastic analysis of Bernoulli processes. *Probability Surveys* **5**, 435-483. MR2476738 MR2476738
- [35] N. Privault and W. Schoutens (2002). Discrete chaotic calculus and covariance identities. *Stoch. Stoch. Reports* **72**, 289-315. MR1897919 MR1897919
- [36] G. Reinert (2005). Three general approaches to Stein's method. In: *An introduction to Stein's method*, 183-221. Lect. Notes Ser. Inst. Math. Sci. Natl. Univ. Singap. **4**, Singapore Univ. Press, Singapore. MR2235451 MR2235451
- [37] G. Reinert and A. Röllin (2009). Multivariate normal approximation with Stein's method of exchangeable pairs under a general linearity condition. *Ann. Probab.* **37** , 2150-2173. MR2573554 MR2573554
- [38] Y. Rinott and V. Rotar (1996). A multivariate CLT for local dependence with  $n^{-1/2} \log n$  rate and applications to multivariate graph related statistics. *J. Multivariate Anal.* **56**, 333–350. MR1379533 MR1379533
- [39] Y. Rinott and V. Rotar (1997). On coupling constructions and rates in the CLT for dependent summands with applications to the antivoter model and weighted  $U$ -statistics. *Ann. Appl. Probab.* **7**, 1080–1105. MR1484798 MR1484798
- [40] V.I. Rotar' (1975). Limit theorems for multilinear forms and quasipolynomial functions. *Teor. Veroyatnost. i Primenen.*, **20**(3), 527-546. MR0385980 MR0385980
- [41] V.I. Rotar' (1979). Limit theorems for polylinear forms. *J. Multivariate Anal.* **9**, 511–530. MR0556909 MR0556909
- [42] R.P. Stanley (1997). *Enumerative Combinatorics, Vol. 1*. Cambridge University Press. Cambridge. MR1442260 MR1442260

- [43] C. Stein (1972). A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. In: *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, Vol. II: Probability theory*, 583–602. Univ. California Press, Berkeley, CA. MR0402873 MR0402873
- [44] C. Stein (1986). *Approximate computation of expectations*. Institute of Mathematical Statistics Lecture Notes - Monograph Series, 7. Institute of Mathematical Statistics, Hayward, CA. MR0882007 MR0882007
- [45] D. Surgailis (2003). CLTs for Polynomials of Linear Sequences: Diagram Formulae with Applications. In: *Long Range Dependence*. Birkhäuser, Basel, 111-128. MR1956046 MR1956046
- [46] W.R. van Zwet (1984). A Berry-Esseen bound for symmetric statistics. *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **66**, 425-440. MR0751580 (86h:60063) MR0751580

## Appendix: Some technical proofs

### Proof of Lemma 2.4

In what follows, for  $k \geq 1$ , we shall write  $\mathbf{a}_k$  to indicate the generic vector  $\mathbf{a}_k = (a_1, \dots, a_k) \in \mathbb{N}^k$ . We start by proving the first point. To do this, just write, using the Cauchy-Schwarz inequality,

$$\begin{aligned}
\|f \star_r^l g\|_{\ell^2(\mathbb{N})^{\otimes(n+m-r-l)}}^2 &= \sum_{\mathbf{i}_{m+n-r-l}} f \star_r^l g(\mathbf{i}_{m+n-r-l})^2 \\
&= \sum_{\mathbf{i}_{n-r}} \sum_{\mathbf{j}_{m-r}} \sum_{\mathbf{k}_{r-l}} \left( \sum_{\mathbf{a}_l} f(\mathbf{i}_{n-r}, \mathbf{k}_{r-l}, \mathbf{a}_l) g(\mathbf{j}_{m-r}, \mathbf{k}_{r-l}, \mathbf{a}_l) \right)^2 \\
&\leq \sum_{\mathbf{i}_{n-r}} \sum_{\mathbf{k}_{r-l}} \sum_{\mathbf{a}_l} f^2(\mathbf{i}_{n-r}, \mathbf{k}_{r-l}, \mathbf{a}_l) \sum_{\mathbf{j}_{m-r}} \sum_{\mathbf{b}_l} g^2(\mathbf{j}_{m-r}, \mathbf{k}_{r-l}, \mathbf{b}_l) \\
&\leq \sum_{\mathbf{i}_{n-r}} \sum_{\mathbf{k}_{r-l}} \sum_{\mathbf{a}_l} f^2(\mathbf{i}_{n-r}, \mathbf{k}_{r-l}, \mathbf{a}_l) \sum_{\mathbf{j}_{m-r}} \sum_{\mathbf{l}_{r-l}} \sum_{\mathbf{b}_l} g^2(\mathbf{j}_{m-r}, \mathbf{l}_{r-l}, \mathbf{b}_l) \\
&= \|f\|_{\ell^2(\mathbb{N})^{\otimes n}}^2 \|g\|_{\ell^2(\mathbb{N})^{\otimes m}}^2.
\end{aligned}$$

The first part of the second point is obtained by writing

$$\begin{aligned}
\max_j \left( \sum_{\mathbf{b}_{n-1}} f^2(j, \mathbf{b}_{n-1}) \right)^2 &\leq \sum_j \left( \sum_{\mathbf{b}_{n-1}} f^2(j, \mathbf{b}_{n-1}) \right)^2 \\
&= \|f \star_n^{n-1} f\|_{\ell^2(\mathbb{N})}^2 \\
&\leq \max_j \sum_{\mathbf{b}_{n-1}} f^2(j, \mathbf{b}_{n-1}) \times \sum_{j'} \sum_{\mathbf{b}'_{n-1}} f^2(j', \mathbf{b}'_{n-1}) \\
&= \max_j \sum_{\mathbf{b}_{n-1}} f^2(j, \mathbf{b}_{n-1}) \times \|f\|_{\ell^2(\mathbb{N})^{\otimes n}}^2.
\end{aligned}$$

For the second part, we have, by applying (in order) the Cauchy-Schwarz inequality, as well as the previous estimate,

$$\begin{aligned}
& \|f \star_l^{l-1} g\|_{\ell^2(\mathbb{N})^{\otimes(n+m-2l+1)}}^2 \\
&= \sum_j \sum_{\mathbf{k}_{n-l}} \sum_{\mathbf{l}_{m-l}} \left( \sum_{\mathbf{i}_{l-1}} f(j, \mathbf{i}_{l-1}, \mathbf{k}_{n-l}) g(j, \mathbf{i}_{l-1}, \mathbf{l}_{m-l}) \right)^2 = \sum_j \|f(j, \cdot) \star_{l-1}^{l-1} g(j, \cdot)\|_{\ell^2(\mathbb{N})^{\otimes(n+m-2l)}}^2 \\
&\leq \sum_j \sum_{\mathbf{k}_{n-l}} \sum_{\mathbf{i}_{l-1}} f^2(j, \mathbf{i}_{l-1}, \mathbf{k}_{n-l}) \sum_{\mathbf{l}_{m-l}} \sum_{\mathbf{j}_{l-1}} g^2(j, \mathbf{j}_{l-1}, \mathbf{l}_{m-l}) \\
&\leq \max_j \sum_{\mathbf{b}_{n-1}} f^2(j, \mathbf{b}_{n-1}) \times \sum_{\mathbf{i}_m} g^2(\mathbf{i}_m) = \max_j \sum_{\mathbf{b}_{n-1}} f^2(j, \mathbf{b}_{n-1}) \times \|g\|_{\ell^2(\mathbb{N})^{\otimes m}}^2 \\
&= \sqrt{\max_j \left( \sum_{\mathbf{b}_{n-1}} f^2(j, \mathbf{b}_{n-1}) \right)^2} \times \|g\|_{\ell^2(\mathbb{N})^{\otimes n}}^2 \\
&\leq \sqrt{\sum_j \left( \sum_{\mathbf{b}_{n-1}} f^2(j, \mathbf{b}_{n-1}) \right)^2} \times \|g\|_{\ell^2(\mathbb{N})^{\otimes m}}^2 = \|f \star_{n-1}^{n-1} f\|_{\ell^2(\mathbb{N})} \times \|g\|_{\ell^2(\mathbb{N})^{\otimes m}}^2.
\end{aligned}$$

Finally, for the third part, just observe that

$$\|f \star_1^0 f\|_{\ell^2(\mathbb{N})^{\otimes(2n-1)}}^2 = \sum_j \sum_{\mathbf{i}_{n-1}} \sum_{\mathbf{j}_{n-1}} f(j, \mathbf{i}_{n-1}) f(j, \mathbf{j}_{n-1}) = \sum_j \left( \sum_{\mathbf{i}_{n-1}} f(j, \mathbf{i}_{n-1}) \right)^2 = \|f \star_n^{n-1} f\|_{\ell^2(\mathbb{N})}^2$$

and, for  $2 \leq l \leq n$ :

$$\begin{aligned}
& \|f \star_l^{l-1} f\|_{\ell^2(\mathbb{N})^{\otimes(2n-2l+1)}}^2 \\
&= \sum_{i=j} \sum_{\mathbf{k}_{n-l}} \sum_{\mathbf{l}_{n-l}} \left( \sum_{\mathbf{i}_{l-1}} f(i, \mathbf{i}_{l-1}, \mathbf{k}_{n-l}) f(j, \mathbf{i}_{l-1}, \mathbf{l}_{n-l}) \right)^2 \\
&\leq \sum_{(\mathbf{a}_{n-l+1}, \mathbf{b}_{n-l+1}) \in \Delta_{2n-2l+2}^c} \left( \sum_{\mathbf{i}_{l-1}} f(\mathbf{i}_{l-1}, \mathbf{a}_{n-l+1}) f(\mathbf{i}_{l-1}, \mathbf{b}_{n-l+1}) \right)^2 \\
&= \|f \star_{l-1}^{l-1} f \times \mathbf{1}_{\Delta_{2n-2l+2}^c}\|_{\ell^2(\mathbb{N})^{\otimes(2n-2l+2)}}^2.
\end{aligned}$$

## Proof of Lemma 2.6

We have, by bilinearity and using the first point of Lemma 2.4,

$$\begin{aligned}
& \|f_k \star_r^r g_k - f \star_r^r g\|_{\ell^2(\mathbb{N})^{\otimes(m+n-2r)}} = \|f_k \star_r^r (g_k - g) + (f_k - f) \star_r^r g\|_{\ell^2(\mathbb{N})^{\otimes(m+n-2r)}} \\
&\leq \|f_k \star_r^r (g_k - g)\|_{\ell^2(\mathbb{N})^{\otimes(m+n-2r)}} + \|(f_k - f) \star_r^r g\|_{\ell^2(\mathbb{N})^{\otimes(m+n-2r)}} \\
&\leq \|f_k\|_{\ell^2(\mathbb{N})^{\otimes n}} \|g_k - g\|_{\ell^2(\mathbb{N})^{\otimes m}} + \|f_k - f\|_{\ell^2(\mathbb{N})^{\otimes n}} \|g\|_{\ell^2(\mathbb{N})^{\otimes m}} \xrightarrow[k \rightarrow \infty]{} 0.
\end{aligned}$$

## Proof of Proposition 2.9

We assume, without loss of generality, that  $m \leq n$ . We first prove (2.11) for functions  $f$  and  $g$  with support, respectively, in  $\Delta_n^N$  and  $\Delta_m^N$ , for some finite  $N$ . One has that

$$J_n(f) J_m(g) = \int_{\Delta_n^N \times \Delta_m^N} \{f \star_0^0 g\} d\mu_X^{\otimes n+m} = \int_{\Delta_n^N \times \Delta_m^N} \{f \otimes g\} d\mu_X^{\otimes n+m}.$$

For  $r = 0, \dots, m$ , we denote by  $\Pi_r(n, m)$  the set of all partitions  $\pi$  of  $\{1, \dots, n+m\}$  composed of

- i) exactly  $r$  blocks of the type  $\{i_1, i_2\}$ , with  $1 \leq i_1 \leq n$  and  $n+1 \leq i_2 \leq n+m$ ,
- ii) exactly  $n+m-2r$  singletons.

For instance: an element of  $\Pi_2(3, 2)$  is the partition

$$\pi = \{\{1, 4\}, \{2, 5\}, \{3\}\}; \quad (6.73)$$

the only element of  $\Pi_0(n, m)$  is the partition  $\pi = \{\{1\}, \{2\}, \dots, \{n+m\}\}$  composed of all singletons; an element of  $\Pi_3(4, 4)$  is the partition

$$\pi = \{\{1, 5\}, \{2, 6\}, \{3, 7\}, \{4\}, \{8\}\}. \quad (6.74)$$

It is easily seen that  $\Pi_r(n, m)$  contains exactly  $r! \binom{n}{r} \binom{m}{r}$  elements (to specify an element of  $\Pi_r(n, m)$ , first select  $r$  elements of  $\{1, \dots, n\}$ , then select  $r$  elements in  $\{n+1, \dots, n+m\}$ , then build a bijection between the two selected  $r$ -sets). For every  $r = 0, \dots, m$  and every  $\pi \in \Pi_r(n, m)$ , we write  $B_{n,m}^N(\pi)$  to denote the subset of  $\{1, \dots, N\}^{n+m}$  given by

$$\{(i_1, \dots, i_n, i_{n+1}, \dots, i_{n+m}) \in \{1, \dots, N\}^{n+m} : i_j = i_k \text{ iff } j \text{ and } k \text{ are in the same block of } \pi\}$$

(note the “if and only if” in the definition). For instance, for  $\pi$  as in (6.73), an element of  $B_{3,2}^3(\pi)$  is  $(1, 2, 3, 1, 2)$ ; for  $\pi$  as in (6.74), an element of  $B_{4,4}^5(\pi)$  is  $(1, 2, 3, 4, 1, 2, 3, 5)$ . The following two facts can be easily checked;

A)

$$\Delta_n^N \times \Delta_m^N = \bigcup_{r=0}^m \bigcup_{\pi \in \Pi_r(n,m)} B_{n,m}^N(\pi),$$

where the unions are disjoint;

B) For every  $r = 0, \dots, m$ , and every  $\pi \in \Pi_r(n, m)$ ,

$$\int_{B_{n,m}^N(\pi)} \{f \star_0^0 g\} d\mu_X^{\otimes n+m} = \int_{\Delta_{n+m-2r}} \{f \star_r^r g\} d\mu_X^{\otimes n+m-2r}.$$

(note that the last expression does not depend on the partition  $\pi$ , but only on the class  $\Pi_r(n, m)$ ).

It follows that

$$\begin{aligned}
J_n(f)J_m(g) &= \int_{\Delta_n^N \times \Delta_m^N} \{f \star_0^0 g\} d\mu_X^{\otimes n+m} \\
&= \sum_{r=0}^m \sum_{\pi^{(r)} \in \Pi_r^*(n,m)} \int_{B_{n,m}^N(\pi^{(r)})} \{f \star_0^0 g\} d\mu_X^{\otimes n+m} \\
&= \sum_{r=0}^m r! \binom{n}{r} \binom{m}{r} \int_{\Delta_{n+m-2r}} \{f \star_r^r g\} d\mu_X^{\otimes n+m} \\
&\stackrel{(*)}{=} \sum_{r=0}^m r! \binom{n}{r} \binom{m}{r} \int_{\Delta_{n+m-2r}} \{\widetilde{f \star_r^r g}\} d\mu_X^{\otimes n+m} \\
&= \sum_{r=0}^m r! \binom{n}{r} \binom{m}{r} J_{n+m-2r} \left[ (\widetilde{f \star_r^r g}) \mathbf{1}_{\Delta_{n+m-2r}} \right],
\end{aligned}$$

which is the desired conclusion. We stress that the equality (\*) has been obtained by using the symmetry of the measure  $\mu_X^{\otimes n+m}$ . The result for general  $f, g$  is deduced by an approximation argument and Lemma 2.6. This concludes the proof.  $\square$